

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: Apr11 22, 2002, 18:19:47 : Search time 116.57 seconds
(without alignments)
2391.648 Million cell updates/sec

Title: US-09-609-146-24

Perfect score: 1231

Sequence: 1 gttgtgatttaagtcag.....aggagtgtcagaagcctc 1231

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PTOS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/Dockfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	133.2	10.8	1092	4	US-09-077-675A-15
2	132.8	10.8	1088	4	US-09-077-675A-6
3	128	10.4	1063	4	US-09-077-675A-1
4	103.4	8.4	836	4	US-09-077-675A-11
5	103.2	8.3	3129	4	US-09-077-675A-14
6	100.2	8.1	1122	4	US-09-077-675A-9
7	98.4	8.0	1029	4	US-09-077-675A-4
8	95.6	7.8	1161	1	US-08-086-439C-2
9	95.6	7.8	1161	1	US-08-434-877-2
10	95.6	7.8	1367	3	US-08-475-742-3
11	95.6	7.8	1370	1	US-08-056-051-1
12	95.6	7.8	1370	1	US-07-928-611-17
13	95.6	7.8	1370	2	US-08-487-811A-17
14	95.6	7.8	1370	4	US-09-060-694-17
15	95.6	7.8	1370	5	PCT-US93-07370-17
16	95.6	7.8	1466	1	US-08-056-051-3
17	95.6	7.8	1466	1	US-07-928-611-19
18	95.6	7.8	1466	2	US-08-487-811A-19
19	95.6	7.8	1466	4	US-09-060-694-19
20	95.6	7.8	1466	5	PCT-US93-07370-19
21	95.6	7.8	1610	1	US-08-056-051-5
22	95.6	7.8	1610	1	US-07-928-611-21
23	95.6	7.8	1610	2	US-08-487-811A-21
24	95.6	7.8	1610	4	US-09-060-694-21
25	95.6	7.8	1610	5	PCT-US93-07370-21
26	91.6	7.4	1529	3	US-08-858-876A-3
27	91.6	7.4	1529	4	US-09-472-880-3

28	79.6	6.5	283	4	US-08-993-088A-4	Sequence 4, Appl1
29	76.8	6.2	1205	1	US-08-417-103-13	Sequence 13, Appl1
30	76.8	6.2	1634	1	US-07-816-283-1	Sequence 1, Appl1
31	76.8	6.2	1634	1	US-08-417-103-1	Sequence 1, Appl1
32	70.2	5.7	1261	3	US-08-475-742-16	Sequence 16, Appl1
33	70.2	5.7	1410	3	US-08-147-592A-1	Sequence 13, Appl1
34	69.4	5.6	936	2	US-08-288-663A-13	Sequence 2, Appl1
35	69.4	5.6	1194	2	US-08-288-663A-2	Sequence 3, Appl1
36	69.4	5.6	1228	2	US-08-288-663A-3	Sequence 1, Appl1
37	68.8	5.6	1575	4	US-08-858-876A-1	Sequence 1, Appl1
38	68.8	5.6	1575	4	US-09-472-880-1	Sequence 1, Appl1
39	68.6	5.6	1342	3	US-08-832-399-1	Sequence 1, Appl1
40	68.6	5.6	1342	4	US-09-372-498-1	Sequence 1, Appl1
41	67.6	5.5	1265	1	US-07-816-283-3	Sequence 3, Appl1
42	67.6	5.5	1265	1	US-08-417-103-3	Sequence 3, Appl1
43	66.8	5.4	1085	2	US-08-466-103A-11	Sequence 11, Appl1
44	66.2	5.4	1164	4	US-08-993-088A-6	Sequence 6, Appl1
45	66.2	5.4	1601	1	US-08-722-001-7	Sequence 7, Appl1

ALIGNMENTS

RESULT 1
US-09-077-675A-15
; Sequence 15, Application US/09077675A
; Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feigner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Peng, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P. O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077, 675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1092 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-077-675A-15
Query Match 10.8%; Score 133.2; DB 4; Length 1092;

Query Match	8.4%	Score 103.4	DB 4	Length 836
Best Local Similarity	49.1%	Pred. No. 1.1e-19		
Matches	362	Conservative	0	Mismatches 365
			Indels	9
			Gaps	3

QY	275	ctggtcctcgtctcttcttggagatgctcttggaaatactacagatgttgacacattacccttc	334
Db	3	CTGCTCATCTTCTCTGATGCTGCCCTGGACCTCGTTCCGCTCGGAGATACCGGGCCG	62
QY	335	ctgttcaggccctgttggaaagtctacttcaagaagccctcttcgagaactgtgtgttcgc	394
Db	63	AACCTTGGGAGCTCTCTGCAAACTCTTCCAAATTCGTAGTAGAGAGCTGACCTACGCG	122
QY	395	lccattctcaagtctacccaacgcttagcgttagagcgctatgttgcaacttc	454
Db	123	ACGGTGCTACCATTCACACAGCGCTGAGCGCTGAGCGCTACTTGGCCATCTGCTTCCACATC	182
QY	455	cgaagccaagcttgagagaagcaagcgcgcgaagcgccttcagaatcccaagctcttgagac	514
Db	183	CGGGCCAAGGTGTGTGTACCAAGGGGGCGGTGAAGCTGTATCTTCTGCTATCGGAGCC	242
QY	515	tctctcttggtcttctcttcttgcccaataacagatccatccatgataagttccagcactt	574
Db	243	GTGGCTTCTGACGCGCCGGGCCCATCTTCTGTGCTAAGTCGGGGGTGGAGACAGAAAGCGC	302

```

0Y 575 cccaacgggtctccgcgtacccgtcgtcgaagcaacccgtgacaagctacacaacccaagtgtggtg 634
0Y 575 cccaacgggtctccgcgtacccgtcgtcgaagcaacccgtgacaagctacacaacccaagtgtggtg 634
Db 303 ACCGACCCCTTGGGACACCAACGAGtGCCGCCCCACCGAGTTTGGGtGCCCTCTGGACtGG 362
0Y 635 tataactgcatcaccacgaagccagctccctctctacatccctcccaatgacccctc 694
Db 363 CTCACGGTCATGtGGtGGGTGTCCAGCATCTTCTTCT--TCCTCTCTGTCTTCTGTCTC 419
0Y 695 agcgctccctacatccctcctcgaaggtcgaagctgaagagatgaaatcccttgaggcgac 754
Db 420 ACGGTCCTCTACAGCTCATCTGCGAGAACTGTGGCGAGAGAGACCGCGCGATGCTGTCTC 479
0Y 755 aaagtgagctgtgaaatattcacaagacccctctgaagatgcacccaaagtgcgtgtgtc 814
Db 480 --GGGGGTGCTCGCTGCTAGGAGGACAGAACCAACGAAAGCAAACTGAAATGCTGGCTGA 536
0Y 815 ttggtcccgctgtctgtccatctgtcgtgaacccctccatctgagtcgq---gctctctc 871
Db 537 GTGGGTGTTGGCTTCATCTCTGTGCTGCCCTTCCACGATAGGGCGCATATTATTTTCC 596
0Y 872 agctttgtgaagagatgagacagatcccttgcgtcgtgtgttcaacccatcatatgtgta 931
Db 597 AAATCTTTGAGACCTGTGCTCTCTTGAGATGCTCAGATACACCCAGTCTGCAACCTCTGTG 656
0Y 932 tcaagtgctctcttattactgtagctccgcggtcaaccccatatctatacctctctct 991
Db 657 TCTCTTGTCTCTCTTACTCTAGTGTGCTCATCAACCCCATTTCTGTACAACTATGTCTC 716
0Y 992 cggcgctccgagcgagc 1008
Db 717 AAGAAGTACGGGtGGC 733

RESULT 5
US-09-077-675A-14
Sequence 14, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pond, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H. T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
INFORMATION FOR SEQ ID NO: 14:

```



```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegretti & Witcoff, Ltd.
STREET: 10 South Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/928,611
FILING DATE: 19920810
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: NO. 5569601nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 810-221-8317
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1370 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..103
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 1268..1370
FEATURE:
NAME/KEY: CDS
LOCATION: 104..1267
US-07-928-611-17

```

[illegible]

Db 570 TCGGGCCACGTGGCTGCTGTCTCCCGGGGTGGCGGCCCCGTACTGTGGGGCTCAACG 629

Qy 558 tcaagttccagcactttcccaagggctcctcgt 591

Db 630 ACGTCCGGCGCGGACACCCCGCTGTCCGCCCT 663

```

RESULT 13
US-08-487-811A-17
: Sequence 17, Application US/08487811A
: Patent No. 5883226
: GENERAL INFORMATION:
: APPLICANT: Civeill, Olivier
: APPLICANT: Van Tol, Hubert H.M.
: TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
: STREET: 300 South Wacker Drive
: CITY: Chicago
: STATE: IL
: COUNTRY: USA
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/487,811A
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: No. 5883226nan, Kevin E
: REGISTRATION NUMBER: 35,303
: REFERENCE/DOCKET NUMBER: 90,1092-L
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-913-0001
: TELEFAX: 312-913-0002
: TELEX:
: INFORMATION FOR SEQ ID NO: 17:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1370 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cdna
: FEATURE:
: NAME/KEY: 5'UTR
: LOCATION: 1..103
: FEATURE:
: NAME/KEY: 3'UTR
: LOCATION: 1268..1370
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 104..1267
: US-08-487-811A-17

```

	Query Match	Similarity	7.8%	Score 95.6	DB 2	Length 1370
Best Local Similarity	50.7%			Pred. No. 2	1e-17	
Matches 230	Conservative	0	Mismatches 224	Indels	0	Gaps 0
OY	138	cggcgctcgtgagcctatgcgcgtcatcttcctcgtggtgggtgaatgagcaactctctgtagt	197			
Db	210	CGGCCCTGGTGGGGGGCGTGCCTCATTCGGCGCGGTGCTCCGGGGAACTCGCTGCTT	269			
OY	198	gcacgtgtagtcttcgcagacacagactcttgaagacacccacacactactctcttcagct	257			
Db	270	GGGTAGCGGTGGCCACCGAGCGCGCCCTTCACAGCCGCCACCACTCTTCATCGTAGGCC	329			
OY	238	tgagcagctcagacatctgctgctgctctctcttgaggatgagctcttgaaatcagaagatg	317			

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:17:43 : Search time 25.8 Seconds
(without alignments)
1166.238 Million cell updates/sec

Title: US-09-609-146-25
Perfect score: 2076
Sequence: 1 MGKLENSWHDPLMKYTLNS.....QSSIHNTNLTPACAGEVP 395

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Prctd. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	513	24.7	424	2 JH0164	neurotensin recept
2	499	24.0	418	2 S29506	neurotensin recept
3	457.5	22.0	378	2 T15816	hypothetical prote
4	430.5	20.7	416	2 S68822	neurotensin recept
5	416	20.0	398	2 JN0708	thyrotropin-releas
6	412	19.8	393	2 A39251	thyrotropin-releas
7	409	19.7	402	2 I56595	neurokinin 2 recep
8	403.5	19.4	411	2 I56444	thyrotropin-2 relea
9	403.5	19.4	412	2 S23436	thyrotropin-2 recep
10	402.5	19.4	392	2 S65693	opioid receptor mu
11	400.5	19.3	400	2 I56553	mu opiate receptor
12	399.5	19.2	384	2 S00516	neurokinin 2 recep
13	394	19.0	398	1 J01059	neurokinin 3 recep
14	387.5	18.7	385	2 S55524	neurokinin 3 recep
15	387	18.6	398	2 A57510	mu opioid receptor
16	381.5	18.4	394	2 J07209	galanin receptor -
17	380.5	18.3	452	2 A34916	neurokinin 3 recep
18	380	18.3	398	2 I56517	mu opioid receptor
19	373	18.0	384	2 S20303	neurokinin 2 recep
20	372	17.9	398	2 I56504	mu opioid receptor
21	368	17.7	390	2 A36737	neurokinin 2 recep
22	366.5	17.7	465	1 J01517	neurokinin 3 recep
23	363.5	17.5	384	2 A47249	brain-specific som
24	362.5	17.5	363	2 I57940	somatostatin recep
25	362	17.4	380	2 S36143	kappa opioid recep
26	362	17.4	384	2 I57957	neurokinin 2 recep
27	360	17.3	380	2 A55259	kappa opioid recep
28	359.5	17.3	380	2 JC2338	kappa opioid recep
29	359	17.3	369	2 B41795	somatostatin recep

30	359	17.3	440	2 A44081	kappa-type opioid
31	357	17.2	372	2 S34592	delta opioid recep
32	357	17.2	382	2 A48227	kappa opioid recep
33	357	17.2	504	2 A41783	tachykinin recepto
34	356	17.1	388	2 JN0605	somatostatin recep
35	356	17.1	399	2 S29480	bombesin receptor
36	355	17.1	369	2 J02083	somatostatin recep
37	354.5	17.1	352	2 J02096	thyrotropin releas
38	354.5	17.1	399	2 A46632	bombesin-like pept
39	353.5	17.0	384	2 JC4629	somatostatin recep
40	351.5	16.9	372	2 I38532	delta opioid recep
41	351	16.9	369	2 A45291	somatostatin recep
42	350	16.9	369	2 D41795	somatostatin recep
43	349	16.8	380	2 JC2434	kappa opioid recep
44	347.5	16.7	418	2 A46226	somatostatin recep
45	347	16.7	346	2 S29248	somatostatin recep

ALIGNMENTS

RESULT 1
JH0164
neurotensin receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #extl_change 17-Mar-2000
C:Accession: JH0164
R:Tanaka, K.; Masu, M.; Nakanishi, S.
Neuron 4, 847-854, 1990
A>Title: Structure and functional expression of the cloned rat neurotensin receptor.
A:Reference number: JH0164; M0ID:90297956
A:Accession: JH0164
A:Molecule type: mRNA
A:Residues: 1-424 <TRAN>
C:Comment: Neurotensin receptor belongs to the family of G protein-coupled receptor.
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F:65-87/Domain: transmembrane #status predicted <TM1>
F:97-121/Domain: transmembrane #status predicted <TM2>
F:144-165/Domain: transmembrane #status predicted <TM3>
F:189-210/Domain: transmembrane #status predicted <TM4>
F:236-260/Domain: transmembrane #status predicted <TM5>
F:309-330/Domain: transmembrane #status predicted <TM6>
F:348-372/Domain: transmembrane #status predicted <TM7>
F:4,38,42/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match	24.7%	Score 513:	DB 2:	Length 424:
Best local similarity	31.6%	Pred. No. 1,1e-35:		
Matches 119:	Conservative 76:	Mismatches 113:	Indels 68:	Gaps 13:
OY	19	NSTEEYLHLGPRKSLDP-----VSVAVALIFLVGNGNLLVCWIVRH--QT 67		
DB	42	NTSESDPA--GP-NSDLVDNTDIYSKVLATYALFLVVGTVGNSVTAFTRKSLQS 97		
OY	68	LKPTNYLFLSLVSDLVLLGMPLEIYE-MHNNYFELGPGCC---YKRTALFEYVPC 123		
DB	98	LOSTVHYHLGSLSLDILLLLAMPVELYNFIVNHHMARGDAGCRGYF--LRDCTY 154		
OY	124	ASLISVTVVERVAIVHPRAKLESTRRLALRILSLWSFSVFSLPMTSHIGIKFQH 183		
DB	155	ATPLNVAISVERLYLALCHFKAKTLMRSRRTKFKISALIASALAIPLMLFMGL--QN 212		
OY	184	EPNGSSVPGSATCTYTKPMVNVMLIIQANISFLYILPMTLISVLVYLMGLRKDESLA 243		
DB	213	RSGDGTHPGGLVCTPIVDATVVKVIVQVNTFMSFLFPMVLVISIL-----NTVIA 261		
OY	244	NKVAVINHR-----PBRKVTK---MLFVLVLVAICW 273		
DB	262	NLLIVNVAHQAEQRCVCTVHNGLESHSTFNMTTEPVOALRHGVLAIVAVIAFVWC 321		
OY	274	TPPHVDRLEFFSFV--EEWTESLAAVFNLIHVSGVFPEYLSAVANPIIYNLSRFRRAFR 331		

Db 322 LPYHVRIMLCYISDEQWTFLEDFEYHYFALNLALFYSSALNPILYLSANFRQVFL 381
 QY 332 NVVSPCK-WCHPRHR 346
 Db 382 STLACLCPRGMRHRRKK 397

RESULT 2
 S29506

neurotensin receptor - human
 C:Species: Homo sapiens (man)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-2000
 C:Accession: S29506
 R:Vita, N.; Laurent, P.; Lefort, S.; Chalou, P.; Dumont, X.; Kaghad, M.; Gully, D.; Le F
 FERS Lett. 317, 139-142, 1993
 A:Title: Cloning and expression of a complementary DNA encoding a high affinity human ne
 A:Reference number: S29506; MUID:93154505
 A:Accession: S29506
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-418 <VIT>
 A:Cross-references: EMBL:X70070; NID:g35020; PIDN:CAA49675.1; PID:g35021
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 24.0%; Score 499; DB 2; Length 418;

Best Local Similarity 33.0%; Pred. No. 1,6e-34; Mismatches 103; Indels 68; Gaps 12;

QY 40 VSAVALFLVGVGNLVCVIVR---OTLKPITYFLSLAVSDLVLLGMPLEIY 96
 Db 66 VFAVYALFLVGVGNVTFARFTLARKKSLQSLQSTVHYHGLSLDLILLAMPVELY 125
 QY 97 E-WMHNPFLFGVGC---YFKTALFETVCFASLSTVTVSVERVAIVHPRAKLESTR 152
 Db 126 NFTWVHHPMAFGDAGCGYF---LRDACTYATALNVAASLVERVYLAICHPFAKTLMSR 182
 QY 153 RRALRILSLVWSFVSFSLP-----NTSINGIKFQHPNNGSSVGSATCVTKPMV 204
 Db 183 SRKKKFTSAIMLASLITVMTLMEQNSADG---QH-----AGGLVCTPTHTAT 232
 QY 205 YNLIGATSEFLYLPMTLISLVLYLGLRLKRDSELEANKVAVNIHR----- 252
 Db 233 VKVVIQVNTFMSIFPMVISVL-----NTIANKLTVMVQAQAEQGVCTVGC 281
 QY 253 -----PSR-----KSTKMLFVLVLFALCMTFPHYDRLEFSFV--EETTESLAAY 296
 Db 282 EHSTFSMALEPGRVQALRHGVRVLRRAVIAFVVCWLPYHRRIMFCYISDEQWTFLYDF 341
 QY 297 FNLIHVSGVFYLSAVNPDIINLSRRPRAFRNVSPCK-WCHPRHR 347
 Db 342 YHFEYVNTALFYVSSITINILNLVANSANRHHFLATLACLCPVWRRRRKR 393

RESULT 3

T15816
 hypothetical protein C48C5.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000
 C:Accession: T15816
 R:Favell, A.
 submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans cosmid C48C5.
 A:Reference number: Z18410
 A:Accession: T15816
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-378 <FAV>
 A:Cross-references: EMBL:U39994; NID:g1055102; PID:g1055105; PIDN:AA37017.1; GSPDB:GNOC
 A:Experimental source: strain Bristol N2; clone C48C5
 C:Genetics:

A:Gene: CESP:C48C5.1
 A:Map position: X
 A:Introns: 53/3; 87/2; 136/2; 169/3; 209/3; 231/3; 259/3; 286/1; 327/1
 C:Superfamily: adenostine receptor A1

Query Match 22.0%; Score 457.5; DB 2; Length 378;

Best Local Similarity 29.9%; Pred. No. 4.5e-31; Mismatches 109; Conservative 74; Mismatches 140; Indels 41; Gaps 10;

QY 14 LMKYLNSTBE-----YLAHLCGPKRS--DSL-PYSVAVALI 47
 Db 2 LQACLNTTEQDCCLAFNCPIYVSHSESEKAEYMECHFSKRALDVTLYKYVALYIFI 61
 QY 48 FIVGVGNLVCVIVRHOQKPTNYLYFLSLAVSDLVLLGMPLEIYEMHNPFLFG 107
 Db 62 FLVGVGNNTTTCVLMKRNHPMKTHASMYLNLAVSDLVLCVGLPEFVMMNQYPPFP 121
 QY 108 PVGCYETALFETVCFAISLSTVTVSVERVAIVHP-FRAKLESTRRALRILSLVWSFS 166
 Db 122 DYICNLKALAEFTSSVSILILIFALERVAVACHPLMKVQPKFNICTIIGFTWIFS 181
 QY 167 VFSLEPNTSIHGKE--QHFP---NGSSVPGSATCTVT-----KPMVYLLIQATSEFL 216
 Db 182 ILCPMPAIIHHRADYIKMSWPGTDNRIPVKSCKMCAVAFEPKLAETFKILFHSIAIF 241
 QY 217 YILPMTLSVLYLMLGLRLKRDSELEANKVAVNIHRSRKSVMKMLFVLVLFALCMTF 276
 Db 242 FALPFTVILYRIACKVSSNRITQPE--LDTEELORINAIICAIYSAFFICLUP 299
 QY 277 HYDRLEFSFEVENTESLAAVNLIHVSGVFYLSAVNPDIINLSRRPRAFRNVSP 336
 Db 300 QIQRLLEFFEDN-EVILTWNQWYFISGLFPLATIIINIAVNLASSRRRAFKDILID 358
 QY 337 TCKW 340
 Db 359 YC-W 361

RESULT 4

S68822
 neurotensin receptor 2, levocabastine-sensitive - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 20-Jun-2000
 C:Accession: S68822
 R:Chalon, P.; Vita, N.; Kaghad, M.; Gullienot, M.; Bonnin, J.; Delpech, B.; Le Fur, G
 FEBS Lett. 366, 91-94, 1996
 A:Title: Molecular cloning of a levocabastine-sensitive neurotensin binding site.
 A:Reference number: S68822; MUID:96228041
 A:Accession: S68822
 A:Molecule type: mRNA
 A:Residues: 1-416 <CHA>
 A:Cross-references: GB:X97121; NID:g1483579; PIDN:CAA65787.1; PID:g1483580
 A:Experimental source: hypothalamus
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: G protein-coupled receptor; transmembrane protein
 F:33-58/Domain: transmembrane #status predicted <TM1>
 F:70-91/Domain: transmembrane #status predicted <TM2>
 F:110-131/Domain: transmembrane #status predicted <TM3>
 F:155-175/Domain: transmembrane #status predicted <TM4>
 F:204-230/Domain: transmembrane #status predicted <TM5>
 F:296-315/Domain: transmembrane #status predicted <TM6>
 F:335-361/Domain: transmembrane #status predicted <TM7>

Query Match 20.7%; Score 430.5; DB 2; Length 416;

Best Local Similarity 31.5%; Pred. No. 9.4e-29; Mismatches 113; Conservative 61; Mismatches 126; Indels 59; Gaps 11;

QY 44 YALIFLVGVGNLVCVIVRHOQKPTNYLYFLSLAVSDLVLLGMPLEIYEMHNPFLFG 101
 Db 39 YSLIFAGTGNALSVHYVLKARAGRGRLRYHVLALSALLLVSMPELNYNFWWSH 98

```

QY 102 xPpLEPRGCG---vEKTLfETyCpASLSTyTtVvEXyVvIvIHfRfAKLESTRRARi 158
Db 99 YPMvFGDJGCGRYF---vRELCAIATvLsvASUSLAEBClAVCOPLRARLLTPARTRRl 155
QY 159 LSLvNSFSvFSLPNTSIHGikfQ-HPfNGSSvPSACtCTyTKRPMyvNLIIQASfLEy 217
Db 156 LSLvYvASLGLALRPMAvIMGCKHEvESADGEPEPAsvCTyLVSrATLQvfiQvNvLVSf 215
QY 218 ILPMfLISVL-----YfLNG-----LRLKRDE----- 239
Db 216 ALPLALTLfLNGITvNHLMAlySQvPSASAOvSSIPSLRELSLSEGGfLTWRKtSLG 275
QY 240 ---SLfEAKvAVNIHRPSRKsvTKMLfVLvLFAICtKfPfvHvDLfESfVEE- -WTESLA 294
Db 276 vQASLVRHKDASQIR - -SLQHSAQvLRAIvAVyICwLpYHARRMLCYIIDDGvNTNfLY 333
QY 295 AVfNLIHvSGvFFLSSAvNPIIYNLLSRKfRAfRvNvSPTCKwCH-----PRHRQ 348
Db 334 DfYvHfYvYvNTPLfVSSAvTPILYNvASSfFRKfLESLSLGCEQHSvPLPDvEAE 392

```

RESULT 5
JN0708
thyrotropin-releasing hormone receptor - human
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Jun-2000
C:Accession: S40682; JN0759; S50151; S50152; I38356; JN0708
R:Matre, V.; Karlsen, H.E.; Wright, M.S.; Lundell, I.; Fjeldheim, A.K.; Gabrielsen, O.S.
A:Title: Molecular cloning of a functional human thyrotropin-releasing hormone receptor
A:Reference number: S40682; MUID:93371401
A:Accession: S40682
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-398 <MAT>
A:Cross-references: EMBL:X75071; NID:9404157; PID:CA52965.1; PID:9404158
R:Yamada, M.; Monden, T.; Satoh, T.; Satoh, N.; Murakami, M.; Iritanihama, T.; Kakegawa,
Biochem. Biophys. Res. Commun. 195, 737-745, 1993
A:Title: Pituitary adenomas of patients with acromegaly express thyrotropin-releasing ho
pter gene.
A:Reference number: JN0759; MUID:93384596
A:Accession: JN0759
A:Molecule type: mRNA
A:Residues: 1-398 <RYAM>
A:Cross-references: GB:D16845; NID:9577631; PID:BA04120.1; PID:9577632
R:Hiunuma, S.; Hosoya, M.; Ogi, K.; Tanaka, H.; Nagai, Y.; Onda, H.
Biochim. Biophys. Acta 1219, 251-259, 1994
A:Title: Molecular cloning and functional expression of a human thyrotropin-releasing ho
A:Reference number: S50151; MUID:95002135
A:Accession: S50151
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-263 <HIN>
A:Accession: S50152
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 267-388 <HI2>
R:Duthie, S.M.; Taylor, P.L.; Anderson, L.; Cook, J.; Eldne, K.A.
Mol. Cell. Endocrinol. 95, R11-R15, 1993
A:Title: Cloning and functional characterisation of the human TRH receptor.
A:Reference number: I38356; MUID:94063224
A:Accession: I38356
A:Status: translated from GB/EMBL/DD8J
A:Molecule type: mRNA
A:Residues: 1-398 <RES>
A:Cross-references: EMBL:X72089; NID:9440155; PID:CA50979.1; PID:9440156
C:Genetics:
A:Gene: GDB:TRHR
A:Cross-references: GDB:228955; OMIM:188545
A:Map position: Bq23-8q23
A:Superfamily: adenosine receptor A1
C:Keywords: G protein-coupled receptor; receptor; transmembrane protein
C:29-51;Domain: transmembrane #status predicted <TM1>

F:62-83/Domain:	transmembrane	#status predicted	<TM2>
F:101-121/Domain:	transmembrane	#status predicted	<TM>
F:146-168/Domain:	transmembrane	#status predicted	<TM4>
F:194-215/Domain:	transmembrane	#status predicted	<TM5>
F:262-288/Domain:	transmembrane	#status predicted	<TM6>
F:297-319/Domain:	transmembrane	#status predicted	<TM7>

Query Match	20.0%;	Score 416;	DB 2;	Length 398;
Best Local Similarity	31.5%;	Pred. No. 1.5e-27;		
Matches 114;	Conservative 66;	Mismatches 120;	Indels 62;	Gaps 11;

```

OY 33 RSDSLPVSVAVALLELV-----GVGMNLLVCMAIVRHQTLKTPYNYLFSLAVSDLVL 88
Db 17 RAVVALEVOVYTIILLVLLICGLIGTIGNIMVYLVMKRTNMRTPNYCYLSVLAVDMLV 76
OY 89 -LGMPL---EYEMHMYPELEGPVGCYFKTALFETVCFASILSVTVSVERYVAIVHPF 144
Db 77 AAGLEPNITSDIYGSW-----YGVYGCCITFYQLGLGINASSCSITAFITRYIAICHP 131
OY 145 RAKLESTRRLARILISLWVSFVSLEPNTSIHGKIFQHPFGNGSVPSASICTYTKPKMW 204
Db 132 KAOFLCTFSRRAKIIIFWAFETSLCYMLFPLDLNISTYKDAIVI-----SCGYKISNY 187
OY 205 YNLIIQANSFLFIILPMTLISLVYLMGLRL-----KRDESLANKYAV 248
Db 188 YSPITLMDFGVFYVPMILATYLVGFIARILFLNIPSDPKENSKTKMNDSTHONTPLNV 247
OY 249 NIHR-----PSRKSVTMKLFVLVLFVAICWTPFH-----VDRLFPES--FVEEMTESLA 296
Db 248 NTSNCFNSVTSYRSQVQKMLAVVILFALLMMPYRTLIVVYNSFLSPFGENW----- 300
OY 297 FNLHIVSGVFYFLSSAVNPITTYNLISRRFRAPFNVSPTCKMKCHPRHROGPAQKII 356
Db 301 --FLLFCRICIYLSAINPIYVILNMSOKFRAAFR-----KLCNCKOKPTEKPAVSV 350
OY 357 FL 358
Db 351 AL 352

```

RESULT 6
A39251
thyrotropin-releasing hormone receptor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 11-Jan-2000
C:Accession: A39251
R:Straub, R.E.; Frech, G.C.; Joho, R.H.; Gershengorn, M.C.
Proc. Natl. Acad. Sci. U.S.A. 87, 9514-9518, 1990
A:Title: Expression cloning of a cDNA encoding the mouse pituitary thyrotropin-releasing hormone receptor
A:Reference number: A39251; MUID:91088548
A:Accession: A39251
A:Molecule type: mRNA
A:Residues: 1-393 <STR>
A:Cross-references: GB:M59811; GB:M37490; NID:g202153; PIDN:AAA0480.1; PID:g202154
C:Superfamily: adenosine receptor A1
C:Keywords: G protein-coupled receptor; glycoprotein; membrane protein

[illegible]

Db 135 FLCTFSRAKIIIFVMAFTSIYCMALFLLDLNISTYKNAVV-----SCGKISIRNYSP 190
 Qy 208 IIQATSLFYILPMTLSVLYLM-----GLRLKRDSELEANKVAVNIHR 252
 Db 191 IYLMDFCVFVPMILATLVGLFARLLFLNIPIDSPDKSKMKKNOSIONK-NLNLNA 249
 Qy 253 -----PSRKSVTKMLFVLVFAICWTPFH---VDRLFES--EVEWTESLAAVFN 298
 Db 250 TNRCFNSTVSSRKQVTKMLAVVILFALLMMRYRTLVAVNSFLSSPQENK----- 300
 Qy 299 LIHVSGVFYLLSAVNPITYNLISRRFRAFRNVVSPCKCHPRRPOGPAOKIT-- 356
 Db 301 -FLFCRICIYLNAINPVIYNLMSQKFAFR-----KLCNCKOKPTEKANYVAL 352
 Qy 357 -----FLTECHLVLETE 368
 Db 353 NYSYIKESDRSTLELDTIVTD 374

RESULT 7
 156595
 neurokinin 2 receptor - guinea pig
 C/Species: Cavia porcellus (guinea pig)
 C/Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 20-Apr-2000
 C/Accession: I56595
 R/Author: D. Little, J. Thomas, C. Powell, S. Downey-Jones, M. Graham, A. J. Recept. Reg. 14, 399-421, 1994
 A/Title: Isolation and characterization of neurokinin A receptor cDNAs from guinea-pig 1
 A/Reference number: I56595; MUID:95162423
 A/Accession: I56595
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-402 <RES>
 A/Cross-references: GB:S67653; NID:g913274; PIDN:AAB33553.1; PID:g913275
 C/Superfamily: neurokinin 1 receptor

Query Match 19.7%; Score 409; DB 2; Length 402;
 Best Local Similarity 29.3%; Pred. No. 5.9e-27;
 Matches 105; Conservative 65; Mismatches 144; Indels 44; Gaps 11;

Qy 41 SVATLFLVGVGNLLVCAYVHQTLPNTNYFLSLAVSDLLVLLGMPLEIYEMMH 100
 Db 37 ATATLALVLAVVGNAVTMTIILHQRMRTVNYFIYNLALADCMAFNAFNFVYASH 96
 Qy 101 NYPLFPGVCGYFKTALFETVCFASILSVTVSERYVAIVHPRAKLESTRRRALRLIS 160
 Db 97 NI-WYFGRACVYFONLFPTIAMFVSIYSMTALAIIDRYMALVHPQRLSAFSTKA--VIG 153
 Qy 161 LVMSFVSVFSLPNTSINGIKFOHPNGSSVPGSATCTVT-----KPMVYNLIIOATS 213
 Db 154 GIMVVALALAFPOGFYSTI-----TDEGATKCVVAMPEDSRDSSLILHLV---I 202
 Qy 214 FLVYILPMTLSVLYLMGLRKDESLKANKAVNI-HRPSRKSVTKMLFVLVLAIC 272
 Db 203 VLYILPMTLSVLYLMGLRKDESLKANKAVNI-HRPSRKSVTKMLFVLVLAIC 262
 Qy 273 WTPRHVDRLEFSPFEEMTESLAANFNLIHVSGVFYL--SSAVNPITYNLISRRRAA 329
 Db 263 WLPRLHLFLIGSFQED-----YCHKRFIOQYLLAEFLWLAASSTWNPDIYCCLLRRRSG 317
 Qy 330 FRNVVSPCKCHPR-----RHRPOGPAOKITIFLTECHLVLETDAGPOFGOSS 379
 Db 318 FR-LAFRCPCWVPTTEDEKLELHTPSPS-----LRVNRCHTKELLMAGDVTSEAT 369

RESULT 8
 156444
 thyrotropin-releasing hormone receptor - mouse
 C/Species: Mus sp. (mouse)
 C/Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 11-Jan-2000
 C/Accession: I56444
 R/Seller, R.E.: Taylor, P.L.; Lamb, R.F.; Zabavnik, J.; Anderson, L.; Eldne, K.A.

J. Mol. Endocrinol. 10, 199-206, 1993
 A/Title: Functional expression and molecular characterization of the thyrotropin-rel
 A/Reference number: I56444; MUID:93249585
 A/Accession: I56444
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-411 <RES>
 A/Cross-references: GB:S60053; NID:g300151; PIDN:AAB26491.1; PID:g300152
 C/Superfamily: adenosine receptor A1

Query Match 19.4%; Score 403.5; DB 2; Length 411;
 Best Local Similarity 31.3%; Pred. No. 1.7e-26;
 Matches 117; Conservative 61; Mismatches 111; Indels 85; Gaps 13;

Qy 18 LNSTEEYLAHLGCKRRDLSIPSVAVL-----IFLV-----GYMGLLCMYVRH 65
 Db 9 LNQTE-----LPQVAVALEYQVVTLLVAVIGLGIYGNIMVVLVVRT 53
 Qy 66 QTLKPTNYFLSLAVSDLLVL-LGMPL--EYEMMHVYPLFGVPGCYFKTALPETV 121
 Db 54 KHMRTATNCLVSLAVADMLVAAGLPNTDSTYSGW-----YGVYGLCTITLYDYL 108
 Qy 122 CFASILSVTVSERYVAIVHPRAKLESTRRRALRLISLVMSFVSVSLPNTSINGIKF 181
 Db 109 INSSCSITAFETIRYIAICHPKAQFLCTFSRAKIIIFVMAFTSIYCMALFLLDLNI 168
 Qy 182 QHPNGSSVCGSATCTYTKPMVYNLIIOATSLFYILPMTLSVLYLMGLR----- 235
 Db 169 STYKDAIV-----SCGKISIRNYSPILMDFGVYVPMILATVLFARLLFLNP 224
 Qy 236 -----KRDESLANKVAVNIHR-----PSRKSVTKMLFVLVFAICWTPFH- 277
 Db 225 SDRPENSRTKKNDSYHKNKNLNTTNRCSNYSYSSKQVTKMLAVVILFALLMMYRT 264
 Qy 278 --VDRLFES--EVEWTESLAANFNLIHVSGVFYLLSAVNPITYNLISRRRAAFRN 333
 Db 285 LVVAVNSFLSSPQENK-----FLFCRICIYLNAINPVIYNLMSQKFAFR-- 332
 Qy 334 VSPCKCHPRHR 347
 Db 333 -----KLCNCKOKP 341

RESULT 9
 S23436
 thyrotropin receptor - rat
 N/Alternate names: thyrotropin-releasing hormone receptor
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 11-Jan-2000
 C/Accession: S23436; I53279; A49168; P00326
 R/de la Pena, P.; Delgado, L.M.; del Cantino, D.; Barros, F.
 Biochem. J. 284, 891-899, 1992
 A/Title: Cloning and expression of the thyrotropin-releasing hormone receptor from GH
 A/Reference number: S23436; MUID:92322017
 A/Accession: S23436
 A/Molecule type: mRNA
 A/Residues: 1-412 <PEN>
 A/Cross-references: EMBL:X64630; NID:g57394; PIDN:CAA45913.1; PID:g57395
 R/Kimura, N.; Arai, K.; Sahara, Y.; Suzuki, H.; Kimura, N.
 Endocrinology 134, 432-440, 1994
 A/Title: Estradiol transcriptionally and posttranscriptionally up-regulates thyrotrop
 A/Reference number: I53279; MUID:94102223
 A/Accession: I53279
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-412 <RES>
 A/Cross-references: GB:D17469; NID:g464199; PIDN:BA04289.1; PID:g464200
 R/Zhao, D.; Yang, J.; Jones, K.E.; Gerald, C.; Suzuki, Y.; Hogan, P.G.; Chln, W.W.; T
 Endocrinology 130, 3529-3536, 1992
 A/Title: Molecular cloning of a complementary deoxyribonucleic acid encoding the thy
 A/Reference number: A49168; MUID:92283212
 A/Accession: A49168

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-12, 'D', '14-290, 'K', '292-412 <ZHA>
 A:Experimental source: CH cells
 A:Note: Sequence extracted from NCBI backbone (NCBIN:104788, NCBI:P:104795)
 R:Yanada, M.; Monden, T.; Satoh, T.; Iizuka, M.; Murakami, M.; Iriuchijima, T.; Mori, M.
 Biochem. Biophys. Res. Commun. 184, 367-372, 1992
 A:Title: Differential regulation of thyrotropin-releasing hormone receptor mRNA levels
 A:Reference number: PQ0326; MUID:92231953
 A:Accession: PQ0326
 A:Molecule type: mRNA
 A:Residues: 30-58, 'P', '60-222, 'T', '224-261 <YAN>
 A:Experimental source: strain Myster
 A:Note: The authors translated the codon ACA for residue 88 as Ala
 C:Superfamily: adenosine receptor A1
 C:Keywords: G protein-coupled receptor; transmembrane protein
 F:1-22/Domain: transmembrane #status predicted <TM1>
 F:32-54/Domain: transmembrane #status predicted <TM2>
 F:71-92/Domain: transmembrane #status predicted <TM3>
 F:116-140/Domain: transmembrane #status predicted <TM4>
 F:165-186/Domain: transmembrane #status predicted <TM5>

Query Match 19.4%; Score 403.5; DB 2; Length 412;
 Best Local Similarity 31.3%; Pred. No. 1.7e-26;
 Matches 117; Conservative 61; Mismatches 111; Indels 85; Gaps 13;

18 LNSTEEYLAHLGPKRSDLSLVSAVAL-----IFLV-----GVMGNLLVCAYVRAH 65
 9 LNQTE-----LPPQVAVALGXOVVITLLVVICGIGIVGINVAVLVMT 53
 66 QTLKPTNYVYLSLAVSDLVLL-LGMPL---EIEEMHNHYFLPGPGCYEKTALFEIV 121
 54 KIMRATATCYLSLAVADLVAVAGLRITDITIGSW-----VGYGGLCTITLQYIG 108
 122 CFASILSVTVSVERVAIVHPRAKLESTRRRLRLISLVSFSVSLPMTSHIGIKF 181
 109 INASSCSITAFIERHYAICHPIKAQFLCTFSRAKKIIFVWAFISYICMLFPLDLNI 168
 182 QHFPNGSSVPGSATCTVTKPMVNYLLIQATISFLYIILPMTLISVLYIMGRLE----- 235
 169 STYKDAIYI---SCGYKISRNYSPYILMDGVEYVMILATVLYGFIARILEFLNPPI 224
 236 -----KRDESLKANKVAVNIHR-----PSRSKVTMLFLVYVFAICWTPFH- 277
 225 SDPKNSKTKWKDSTHOKNNMLNTNRCFNSTVSSRKQVYMLAVVYLFLLMPPYT 284
 278 --VDRLFPS--FVEEMTESLAVFNLIHVSGVFYLSAVNPDIYINLSRRRAAFRNV 333
 285 LVVNSFLSPQENM-----FLFCRICIYLNAINPIVYINLMSQKFRFAFR-- 332
 334 VSPCKKCHPRHP 347
 333 -----KLCNCKOKP 341

RESULT 10
 S65693
 oploid receptor mu variant MOR1A - human
 C:Species: Homo sapiens (man)
 C:Date: 12-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 05-Nov-1999
 C:Accession: S65693; S51216
 R:Bare, L.A.; Mansson, E.; Yang, D.
 submitted to the EMBL Data Library, July 1994
 A:Description: Expression of two variants of the human mu oploid receptor mRNA in SK-N-SH
 A:Reference number: S65693
 A:Accession: S65693
 A:Molecule type: mRNA
 A:Residues: 1-392 <BAR>
 A:Cross-references: EMBL:U12569; NID:9607911; PIDN:AAB60354.1; PID:9607912
 R:Bare, L.A.; Mansson, E.; Yang, D.
 FEBS Lett. 354, 213-216, 1994
 A:Title: Expression of two variants of the human mu oploid receptor mRNA in SK-N-SH cell

A:Reference number: S51215; MUID:95046336
 A:Accession: S51216
 A:Molecule type: mRNA
 A:Residues: 387-392 <BAN>
 C:Superfamily: vertebrate rhodopsin

Query Match 19.4%; Score 402.5; DB 2; Length 392;
 Best Local Similarity 30.4%; Pred. No. 2e-26;
 Matches 107; Conservative 65; Mismatches 133; Indels 47; Gaps 11;

8 SWIHPDLKMLNSTEEYLAHLGPKRSDL-----SLPVSVALIFLV 50
 30 SWVN-----LSHLDGDLSDPCGPNRTDGGSDLSCLPPTGSPMTAITMALYSIVCV 83
 51 GVMGNLLVCAYVIRHOTLKTPTNYVYLSLAVSDLVLLGMPLE-IYEMHNPPLFGPV 109
 84 GLEGNLVVYVIRYTKMTATNIYIFNLADALATST-LPFGQSVNYLMGTWP--FGYI 140
 110 GCYFPTALFEVYCFASILSVTVSVERVAIVHPRAKLESTRRRLRLISLVSFSVYF 169
 141 LCKIVISIDYVMFYSIFLCTMSVDRIYAVCHPAKALDFTRPAKKIINVCNMLSSAI 200
 170 SLPTSHIGIKFQHPNGSSVPGSATCTVT--KPMVNYLLIQATISFLF-YILPMTLISV 226
 201 GLPVMFIATTKRQ-----GSIDCTITFSPHTWYMNLLKICVFIFAFIMPLVITV 252
 227 LYVGLGRILKRDESLKANKVA-VNIHRPSKSVTKMLFLVYLVFAICWTPFHVDRLFSF 285
 253 CYGLMILRLKSVRLMSGSKEDRNLR---ITRMVLVAVVAVFCWTPPIHIVYIKAL 307
 286 VEWNESLAAVNLIHVSGVFYLSAVNPDIYINLSRRRAAFRNVSP 337
 308 VTIPETTFQTVSWHFCIALG---YTNISCLPVLVAFLEBENFRCFRRCIP 356

RESULT 11
 156553
 mu opiate receptor - human
 N:Alternate names: MOR1 protein; oploid receptor mu
 C:Species: Homo sapiens (man)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 19-May-2000
 C:Accession: 156553; A38991; S41075; S51215
 R:Meatsek, A.; Hurley, J.H.; Bye, L.S.; Campbell, A.D.; Chen, Y.; Tian, M.; Liu, J.; S
 J. Neurosci. 15, 2396-2406, 1995
 A:Title: The human mu opiate receptor: modulation of functional desensitization by ca
 A:Reference number: 156553; MUID:95198115
 A:Accession: 156553
 A:Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-400 <RES>
 A:Cross-references: GB:L29301; NID:9459831; PIDN:AAA73958.1; PID:9459832
 R:Wang, J.B.; Johnson, P.S.; Persico, A.M.; Hawkins, A.L.; Griffin, C.A.; Uhl, G.R.
 submitted to GenBank, August 1994
 A:Reference number: A38991
 A:Accession: A38991
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-50, 'N', '52-233, 'V', '235-400 <MAN>
 A:Cross-references: GB:L25119; NID:9452072; PIDN:AAA20580.1; PID:9452073
 R:Wang, J.B.; Johnson, P.S.; Persico, A.M.; Hawkins, A.L.; Griffin, C.A.; Uhl, G.R.
 FEBS Lett. 338, 217-222, 1994
 A:Title: Human mu opiate receptor. cDNA and genomic clones. pharmacologic characteriz
 A:Reference number: S41075; MUID:94139928
 A:Accession: S41075
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-50, 'N', '52-400 <MAN>
 R:Bare, L.A.; Mansson, E.; Yang, D.
 FEBS Lett. 354, 213-216, 1994
 A:Title: Expression of two variants of the human mu oploid receptor mRNA in SK-N-SH c
 A:Reference number: S51215; MUID:95046336
 A:Accession: S51215

us-09-609-146-25.rpr

```

Query Match      18.7% Score 387.5; DB 2; Length 385;
Best Local Similarity 27.6%; Pred. No. 3.6e-25;
Matches 105; Conservative 64; Mismatches 129; Indels 83; Gaps 11;

Qy      23 EYLALHCGEGRSDLSLPV-----SVAYALIFLVGMGNL 56
       1:-||| | | | | | | | | | | | | | | | | |
Db      32 EWIALDAGNFSSALDLPITTSQAPSOVRNDLTNQFVQPSMRIALMBAVLAVGVAVAVFGNL 91
Qy      57 LYCVATIVRHQTCKPTNYVLFSLAVSDLLY----LLCMLPEIYSMMINYPFLCPVCY 112
       . : . : . : . : . : . : . : . : . : . : . :
Db      92 IYIMIIILAKRRRTYTNNFLVLAISDSAVAFTNLFVNPIYGVSHSEW-----FCANCR 146
Qy      113 EKTALFEYCYCFMSISVTTVSVERVAIVHPRAKLESTRRALRIILSLWSFSVSFELP 172
       . : . : . : . : . : . : . : . : . : . : . :
Db      147 FQNFPPIIAVFASISMTAIAVDRTMAIIDPLKPRLSATKI--VISIWIHLAEFLAF 204

```

Db	205	QCLYSIKV-----MGRTLCYVOMPEGRKQHTHTIIIV---IILVYCFPLDINGVT	253
Oy	228	YLMGLRLKDESEANKVAVNITHRPSRKSVYKMLFLVLVFAICMTPEHVDRLFFSEFE	287

```

Db      254 YTVIGTTLGGEIILDTCDKYEHOQLAKRRKRVKMMIIVVVFPAICMLPRHYHFFITAIYQ 313
OY      288 EMTESLAANFENLIHVSGGEFL-----SSAVNITITNILLSRPRRAFRVNVSPSTWC--- 341
Db      314 QUNR-----WKTIQOYVLASFVLAMSSFTMYNPIIYCCLNKRFRAGFKRAF-----RMCEFI 364
OY      342 -----HPRIRP 347
Db      365 QVSYDELEKTRTHRPTRQRP 385

```

RESULT 15
A57510
mu oploid receptor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 24-Nov-1999
C:Accession: A57510; 148665; 566513; 149300
R:Kautman, D.L.; Keith Jr., D.E.; Anton, B.; Tian, J.; Magendzo, K.; Newman, D.; Tran
J. Biol. Chem. 270, 15877-15883, 1995
A:Title: Characterization of the murine mu oploid receptor gene.
A:Reference number: A57510; MUID:95318184
A:Accession: A57510
A>Status: nucleic acid sequence not shown

A:Cross-references: GB:U19380
R:Min, B.H.; Augustin, L.B.; Felsheim, R.F.; Fuchs, J.A.; Loh, H.H.
Proc. Natl. Acad. Sci. U.S.A. 91, 9081-9085, 1994
A:Title: Genomic structure analysis of promoter sequence of a mouse mu opioid receptor
A:Reference number: 148665; MUID:94377496
A:Accession: 148665
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-398 <RES>
A:Cross-references: EMBL:U10561; NID:9556566; PIDN:AAB60673.1; PID:9565066
R:Rossi, G.C.; Pan, Y.X.; Brown, G.P.; Pasternak, G.W.
FEBS Lett. 369, 192-196, 1995
A:Title: Antisense mapping the MOR-1 opioid receptor: evidence for alternative splicing
A:Reference number: 149300; MUID:95377399
A:Accession: S66513
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-398 <ROS>
A:Cross-references: EMBL:U26915; NID:91055230; PIDN:AA81170.1; PID:91055231
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
C:Genetics:
A:Gene: MOR-1
A:Introns: 95/2, 213/1, 386/3
C:Superfamily: vertebrate rhodopsin

C:keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmembran

Query Match 18.6%; Score 387; DB 2; Length 398;

Best Local Similarity 29.2%; Pred. No. 4.1e-25;

Matches 108; Conservative 66; Mismatches 134; Indels 62; Gaps 13;

```

OY 6 NASMIDPLM-----KYLNSTEEYLAHL-----CGPKRSDL-----36
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9 NISDCSDPLAPASCSPAPGSMLN-----LSHYGNGSDPCGPNRTGLGSHSLCPQTGSP 63
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 37 ----SLPVSVAVALIFLVGVGNLVCMTVIRHQTLKPTNYLFLSLAVSDLVLLGMP 92
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 64 SMYTAITIMALYSIVCVGLGFLVWYIVRYTKMKATATNIYIFNLADALATST-LP 122
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 93 LE-IYEMMHNYPFPGVGCYFETALEFETVCFASILSVTVSERYVAIVFPRAKLEST 151
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 123 FOSVNTLMGTWP--FGNILLCKIVISIDYNNMFTSIFTLCTMSVDRYIACHPVKALDERT 180
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 152 RRRALRLSLVMSFSVFSLPNTSINGIKRQHPNGSSVPSATCTVT--KPMVYNYLII 209
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 PRNAKIYVNCNMILSSAIGLPVFMATTKYRQ-----GSIDCTLTFSHPTWYENLL 232
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 210 QATSELP-YLPMTLISVLYYLMGLRKDESLANKVA-VNIHRSRKSVTKMLEVVL 267
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 233 KICVFTEAFIMPVLIITVCYGLMILRLKSVRLSGSKEDRNLR-----ITRMVLVVA 287
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 268 VFAICWTPFVDRLFESFVEWETESLAAVFNLIHVVGCVFFYLSAVNPILIYNLSRRR 327
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 288 VFIVCWTPRIHIYIIRKALITIPETTFOTYSWHFCIALG---YTNSCUNPVLVAFLDENFK 344
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 328 AAFRNVSPT 337
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 345 RCFREFCIPR 354
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: April 22, 2002, 14:17:44
Job time: 297 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: Apr11 22, 2002, 17:16:21 ; Search time 2258.39 Seconds
(without alignments)
5857.294 Million cell updates/sec

Title: US-09-609-146-24

Perfect score: 1231
Sequence: 1 gcttgatatttaacacag.....aggagtggtcagaagcctc 1231

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estom:*
5: em_estcpl:*
6: em_estlba:*
7: em_estlro:*
8: em_estlov:*
9: em_hlc:*
10: gb_estl:*
11: gb_estc2:*
12: gb_estc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_fod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	281.8	22.9	555	10	BE751626 203809 MA
2	214.2	17.4	465	10	BE751310 203132 MA
3	187.8	15.3	941	13	CNS036WH
4	179.2	14.6	929	11	BE871983 60148032
5	139	11.3	414	11	M45474
6	114	9.3	223	6	BB587450
7	96.6	7.8	1079	13	CNS04P4T
8	86.4	7.0	389	10	AA562357
9	86.4	7.0	450	11	R13850
10	86.4	7.0	500	11	BF603623
11	85.8	7.0	959	13	CNS04GXP
12	83.6	6.8	448	11	BF323227

C 13	82	6.7	480	13	A2616812
C 14	77.8	6.3	379	11	H11359
C 15	77.8	6.3	817	13	CNS01X12
C 16	74.6	6.1	835	13	CNS020N1
C 17	72.2	5.9	471	10	AU204159
C 18	70.8	5.8	741	11	B1224313
C 19	69	5.6	1189	12	AK005368
C 20	68.8	5.6	634	11	BG714306
C 21	67.4	5.5	975	13	CNS039RV
C 22	67.4	5.5	1041	11	BG293064
C 23	66	5.4	612	11	BF446062
C 24	66	5.4	825	10	AU079556
C 25	63.6	5.2	396	10	BE654108
C 26	62.8	5.1	979	13	CNS04BAL
C 27	62.6	5.1	875	10	AL537553
C 28	62.4	5.1	563	13	A2357931
C 29	62	5.0	460	10	AW206493
C 30	61.2	5.0	662	10	A1820631
C 31	61	5.0	788	10	AL541044
C 32	60.4	4.9	462	10	AM138148
C 33	60.4	4.9	544	10	AT344017
C 34	60.4	4.9	549	11	BE938149
C 35	60.2	4.9	638	10	AM139137
C 36	60	4.9	566	10	BE266564
C 37	60	4.9	988	13	CNS018EH
C 38	59.6	4.8	335	13	BH034857
C 39	59.4	4.8	753	10	AT692402
C 40	59.2	4.8	768	10	BE746231
C 41	59	4.8	523	11	BE437994
C 42	59	4.8	861	11	BE798784
C 43	58	4.7	339	13	AO015065
C 44	58	4.7	519	13	AO019411
C 45	58	4.7	597	10	AT693287

ALIGNMENTS

RESULT 1
LOCUS BE751626 555 bp mRNA
DEFINITION 203809 MARC 2B0V Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE751626
VERSION BE751626.1 GI:10165618
KEYWORDS
SOURCE
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 555)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,
G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
Petea,G., Holt,I., Karamecheva,S., Liang,F., Quackenbush,J. and
Keefe,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross-match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACACATGACACAT
BACKWARD: GTTTCACATGACGACG
Plate: 44 row: O column: 6

DB	Accession	Source	Organism	Reference	Authors	Title	Journal	Comment
DB	363	AGAGCCACCAACCA--TACTACCTCTTCAGCGCTGGCCGGGTCTCTGACCTCTCTGGGCTTCTGC	421					
OY	285	tcttgggaatgcctcttgaaatcct---acagatcttggaacaataacccttccctgttcg	341					
DB	422	TCTTTTGGAAAGCCCGCTGGAGGTCTTAATGACATGTGTGGCGCAACTACCTTTTGGGTC	481					
OY	342	gg---ccgttggaatgcctctcaagagacacccctctctgagatctgtgtcttgcctca	398					
DB	482	GGGCGCGTGGGGCTGTACATTCACAGAGGGCCCTTTTTCAGACCGGTGCTTCGCGCTCA	541					
OY	399	tctctagtcctcaacagtgtagcgttaagcgtctatgtgtgcatttccaccccttcagag	458					
DB	542	TCTCTAGATATCACACCG--TCAAGTGTAGGCGCTTACGTGGCAATCTCTACACCCGTT--CGGG	598					
OY	459	ccaagcttgagagacagcgcgcgcgcgcctcaagatctctcagcctagtcgttgagctct	518					
DB	599	CCAAATTGAGGAGACACCGCGGTGTCCGGCGCTCAGAGATCTCGGAATTTCTGTGGGCTACT	658					
OY	519	ctgtgtctctctctctcttgcctcaataccagatctcaatgtgcatca	560					
DB	659	CCGACACTTCTCCCTG--CAAAACCCAGCATCATGTGGCTTCA	698					
RESULT	5							
LOCUS	N45474	414 bp	mRNA	EST	13-FEB-1996			
DEFINITION	yy59b04.r1 Soares_multiple_sclerosis_2NBMHSP	Homo sapiens	CDNA					
ACCESSION	N45474							
KEYWORDS	N45474.1	GI:1186640						
SOURCE	human.							
ORGANISM	Homo sapiens							
REFERENCE	1 (bases 1 to 414)							
AUTHORS	Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rittin, L., Rohlfing, T., Soares, M., Tan, F., Trevisakis, E., Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.							
	The WashU-Merck EST Project							
	Unpublished (1995)							
	Contact: Wilson RK							
	Washington University School of Medicine							
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108							
	Tel: 314 286 1800							
	Fax: 314 286 1810							
	Email: estewatson.wustl.edu							
	This clone is available royalty-free through LNL; contact the							
	IMAGE Consortium (info@image.llnl.gov) for further information.							
	Seq primer: T7							
	High quality sequence stop: 264.							
FEATURES								
SOURCE								
	Location/Qualifiers							
	1..414							
	/organism="Homo sapiens"							
	/db_xref="GDB:3896183"							
	/db_xref="taxon:9606"							
	/clone="IMAGE:277807"							
	/clone_lib="Soares_multiple_sclerosis_2NBMHSP"							
	/sex="male"							
	/tissue_type="multiple sclerosis lesions"							
	/dev_stage="Age 46"							
	/lab_host="DH10B (ampicillin resistant)"							
	/note="Vector: pRT73D (Pharmacia) with a modified							
	polylinker V type; phagemid; Site 1: Not I; Site 2: Eco RI							
	; 1st strand cDNA was primed with a Not I - oligo(dT)							
	primer 15'							
	TGTTTCAATCTGAAGTGGAGCGGCCGCAATTTTTTTTTTTTTTTT 3'							
	double-stranded cDNA was size selected, ligated to Eco RI							
	adapters (Pharmacia), digested with Not I and cloned into							
	the Not I and Eco RI sites of a modified pRT73 vector							
	(Pharmacia). Library went through one round of							
	normalization to a Cot = 5. Library constructed by Bento							

BASE COUNT	113 a	123 c	71 g	105 t	2 others
Query Match					
Best Local Similarity					
Matches 178; Conservative					
0; Mismatches 65; Indels 0; Gaps 0;					
969	ccatcatctatctaccctccctgctctgcggcgcttcgcggcgccgcttcgaatgctctcc	11.3%; Score 139; DB 11; Length 414;			
1	ccattatctatnaactactgctctcccgcttcacagcagcattccagatgcatctctt	73.3%; Pred. No. 2.7e-26;			
1029	ctactctgcaaatgctgcctccctgcgcgcacgcagaggaacctccagcccgaaagatca				
61	ctttccacacaaagatggcgactctccagcattgaccacacagtgctgcacccgagcagca				
1089	tctctctgcagagaatgctcaactcctctgtagcgctgacagagatgacggcccccagtcctg				
121	tctctctgcacacaaagcactcttctgtggagctgacccggaagatattagctcccat				
1149	gtcagctatccatccacaaacacacacacttaccagcgcccccctgtagcagagaggtacct				
181	gtcagctatcattcgcacacacactctcactctccacacagccctctctgtagcagagatgcna				
QY	1209 aaa 1211				
DB	241 GAA 243				
RESULT	6				
LOCUS	BB587450	223 bp	mRNA	EST	30-NOV-2000
DEFINITION	BB587450 RIKEN full-length cDNA clone 993000G13 5' mRNA sequence.				
ACCESSION	BB587450				
VERSION	BB587450.1	GI:11483994			
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
	1 (bases 1 to 223)				
	Alizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T., Carninci,P.				
	Imotani,K., Ishii,Y., Itoh,M., Iizawa,M., Kawai,J., Kojima,Y., Konno				
	H., Kusakabe,M., Matsuyama,T., Miyazaki,A., Nakamura,M., Nishi,K.,				
	Nomura,K., Numazaki,R., Okazaki,T., Okido,T., Owa,C., Sakai,C.,				
	Sakai,K., Sasaki,D., Sato,K., Shibata,K., Shibata,Y., Shingawa,A.,				
	Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Tanaka				
	T., Toyota,T., Watabiki,A., Yamamura,T., Yasunishi,A., Yoshida,K.,				
	Yoshiki,A., Muramatsu,M. and Hayashizaki,Y.				
	RIKEN Mouse ESTs (Alizawa,K. et al. 2000)				
	Unpublished (2000)				
TITLE	Contact: Yoshihide Hayashizaki				
JOURNAL	Laboratory for Genome Exploration Research Group, RIKEN Genomic				
COMMENT	Sciences Center(GSC), Yokohama Institute				
	The Institute of Physical and Chemical Research (RIKEN)				
	1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 220-0045, Japan				
	Tel: 81-45-503-9222				
	Fax: 81-45-503-9216				
	Email: genome-res@gsc.riken.go.jp,				
	URL:http://genome.gsc.riken.go.jp/				
	Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoke,S., Sasakhi				
	,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.				
	Thermotabilization and thermocactivation of thermostable enzymes by				
	thiazole and its application for the synthesis of full length				
	cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)				
	Itoh,M., Katsunai,T., Akiyama,J., Shibata,K., Iizawa,M., Kawai,J.,				
	Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki				
	,Y. and Hayashizaki,Y.				

AUTHORS ROESC-CROLLINS, H., J

TITLE The WashU-HHMI Mouse EST Project

TITLE	Keele, J.W.
JOURNAL	Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
MEDLINE	Genome Res. 11 (4), 626-630 (2001)
COMMENT	21180013
	Contact: Smith TPL

Email: smith@eemail.marc.usda.gov
 single pass sequencing. Bases called and alt-trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -mnscore 18
 and -mismatch 12 options.

FEATURES
FORWARD: AGGAAACAGCTATGACCACT
BACKWARD: GTTTTCCCAGTCACGACG
plate: 49 **row:** I **column:** 6
Seq primer: ATTTAGGTCACACTTAG
Location/Qualifiers

```

source
1..500
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_1lb="MARC 380V"
/tissue_type="pooled"
/lab_host="DH10B"
/notice="Vector: PCMV SPORT6; Site.1: XbaI; Site.2: XhoI;
Library made from pooled tissue from marrow, alveolar
macrophage, ovary, fetal semitendinosus muscle, and fetal
longissimus muscle."
BASE COUNT
71 a 179 c 139 g 111 t
ORIGIN

```

Query Match	7.08;	Score 86.4;	DB 11,	Length 500;
Best Local Similarly	52.88;	Pred. No. 2.7e-12;		
Matches 211; Conservative	0;	Mismatches 186;	Indels 3;	Gaps 1,

Oy	140	gtgtctgttgagctctatgtgcgtgattcttcctctgtgtgggttaatgtagcaattcttcgtgtgc	199
Db	68	gttccacagcactcttcgtggccctcttctgtgtgtgggcatgcgggcaaacctgtctacacatg	127
Oy	200	atgtgtatgtgcgacatcagactttgaagacaccacacacactactctcttcacagt	259
Db	128	ctgtgtgtgtgtgccttttccggagacttcgtaaccacacacacactcttaccttcacagatg	187
Oy	260	gcagatctcagaatctgtctgtgtctctctctgtgggaatgctctgtgaatatcagaagatgtg	319
Db	188	gcctttctccgacttactgacatcttctcctt---gcattgcccttcgacactcgtggccgcctctgg	244
Oy	320	cacaaatacccttctccctgttcctgggacctgtgtagtctactcaagacagacctcttcag	379
Db	245	catttaccggcccttgaaacttggggcgacctctctgcacaaactcttccagttgttcacggag	304
Oy	380	actgtgtgcttctgtccatctcagtcagtcgacacaggttcagctagaagcgtatgtggcc	439
Db	305	agctgcacacttaccgttagcgttaccatcacccgctgagacgttcgacgaccttaccttgcgc	364
Oy	440	atgtccaccccttctcgaagccaaagcttgaagacacgcgtgcagcagggagccctcaagatcttc	499
Db	365	attctgctttcccgctggggcccaagtggtgattcaccaaggccggcggtgaagctgctatc	424
Oy	500	agctagatctgtagactctctctgtgtcttcttcttgcga	539
Db	425	ctggctcatctggccctgtggctttttgtgagcccgccggccca	464

RESULT	11
CNS04GXP	
LOCUS	CNS04GXP 959 bp DNA
DEFINITION	Tetradon nigroviridis genome survey sequence PIC-ori end of clone 10K06 of library G from Tetradon nigroviridis, genomic survey sequence.

ACCESSION	AL290230
VERSION	AL290230.1
KEYWORDS	GI:8028810
SOURCE	GSS: genome survey sequence.
ORGANISM	Tetraodon nigroviridis.
	Tetraodon nigroviridis
	Euryptera; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE	1 (bases 1 to 959)
AUTHORS	Roeest-Crollius,H., Jalllon,O., Dasilva,C., Fitzames,C., Fisher,C., Bouneau,T., Billault,A., Quetler,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 959)
AUTHORS	Roeest-Crollius,H., Jalllon,O., Dasilva,C., Bouneau,T., Fisher,C., Bernot,A., Fitzames,C., Wlnccker,P., Brottlcr,P., Quetler,F., Saurin,W. and Weissenbach,J.
TITLE	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 959)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (12-APR-2000) to the EMBL/Genbank/DBD databases
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon .

FEATURES	SOURCE	Location/Qualifiers
	1. 959	
		/organism="Heterodon nigriviridis"
		/db_xref="taxon:99883"
		/clone="109K06"
		/clone_1b="c"
		/note="Genoscore sequence ID : C0BC109BF03Sp1-end : puc-01"
BASE COUNT	184 a	293 c 290 g 184 t 8 others
ORIGIN		

Query Match	7.0%;	Score 85.8;	DB 13;	Length 559;
Best Local Similarity	55.48;	Pred. NO. 4.8e-12;		
Matches 181; Conservative	3;	Mismatches 140;	Indels 3;	Gaps 1.

OY	157	gctatctctccttgtagggggaatgggcaatctctgtgtgcatgtgaatgtgcgaca	216
Db	408	gcycacgggaatcctgtggggaacggcgtraccatcagggtagccaggtctgaagcggaa	467
OY	217	tcaagacttgaagacaccaccaaatactatctctcagcttggcagltcgaagctgc	276
Db	468	cggtacctggacagaaagacgctacacggacacacatattacagcttgccctgcgcacctcct	527
OY	277	gtccctgctctctggagatgcctctcgaatactc---cgagatgtgycacataaccttc	333
Db	528	gctgctgctcattcgcgcaatcgccgggtgagactctacacggccattcgttccctctacactc	587
OY	334	ccgtgttgagcctcttgagatgctctctctcgaagacagccctctcgaagctgtgtgcttc	393
Db	588	ggcctcgccggcaacgctctcctgttaagamctacaaactctctgttcacagcgcctcacctacgc	647
OY	394	ctccatctctcagtgatccacacggttagcgttagagcgtatgtgacatgtgccacccctc	453
Db	648	cacactttgtgaacattagccacgctcttaccttcgcacggcagctgcccacttgcacccctt	707

```

Qy      454  ccgagcccaagctgtgagagcacgcgagcg  480
          |||  |||  |  |||  |||
Db      708  ccgctacaaaggccctggggcgggaacg  734
          |||  |||  |  |||  |||
RESULT  12

```

```

BF323227      448 bp      mRNA      EST      21-NOV-2000
LOCUS      maa38a07.v1 NCI_CGAP_Brn63 Mus musculus cDNA clone IMAGE:3813037 5'
DEFINITION      similar to SW:NTNR_RAT 063384 NEUROTENSIN RECEPTOR TYPE 2 ;, mRNA
sequence.
ACCESSION      BF323227.1 GI:11272778
VERSION      BF323227
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus.
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 448)
AUTHORS      NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute / National Institute of Neurological
               Disorders and Stroke, Brain Tumor Genome Anatomy Project
               (CGAP/BRGAP), Tumor Gene Index
               Unpublished (1998)
               Other ESTs: maa38a07.x1
               Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-remail.nih.gov
               Tissue Procurement: Jeffrey E. Green, M.D.
               cDNA Library Preparation: Life Technologies, Inc.
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Washington University Genome Sequencing Center
               Clone distribution: NCI-CGAP clone distribution Information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov

MGI:1455149
Seq primer: -40RP from GIBCO
High quality sequence stop: 416.
FEATURES
    source
        1..448
        /organism="Mus musculus"
        /db_xref="taxon:10090"
        /clone="IMAGE:3813037"
        /clone_1lb="NCI_CGAP_Brn63"
        /sex="female"
        /dev_stage="10 weeks"
        /lab_host="DH10B (T1 phage-resistant)"
        /note="Organ: Brain; Vector: PCMV-SPOK6; Site_1: NotI;
        Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
        Average insert size 1.8 kb. Library constructed by Life
        Technologies."
BASE COUNT      53 a      157 c      148 g      89 t      1 others
ORIGIN
Query Match      6.8%; Score 83.6; DB 11; Length 448;
Best Local Similarity 55.5%; Pred. No. 1.4e-11;
Matches 183; Conservative 0; Mismatches 144; Indels 3; Gaps 1;

```

```

Oy      538      caatacagcatcattgcatcaagtcca 567
Db      355      CATGCCGCGTATCATCGGACAGACAGCACA 364

RESULT 13
LOCUS      A2616812/C
DEFINITION      A2616812      480 bp      DNA      GSS      13-DEC-2000
               1M0446J15R Mouse 10kb plasmid UGCCIM library Mus musculus genomic
               clone UGCCIM0446J15 R, DNA sequence.
ACCESSION      A2616812
VERSION      A2616812.1 GI:11739002
KEYWORDS      GSS.
SOURCE      house mouse.
ORGANISM      Mus musculus.
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 480)
AUTHORS      Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
               Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
               M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
               and Wright,D.,Weiss,R.
TITLE      Mouse whole genome scaffolding with paired end reads from 10kb
               plasmid inserts
JOURNAL      Unpublished (2000)
COMMENT      Contact: Robert B. Weiss
               University of Utah Genome Center
               University of Utah
               Km. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
               84112, USA
               Tel: 801 585 5606
               Fax: 801 585 7177
               Email: ddunn@genetics.utah.edu
               Insert length: 10000 Std Error: 0.00
               Plate: 0446 Row: J Column: 15
               Seq primer: CACACAGAAACAGCTATGACC
               Class: plasmid ends
               High quality sequence stop: 480.
FEATURES
    source
        1..480
        /organism="Mus musculus"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="UGCCIM0446J15"
        /clone_1lb="Mouse 10kb plasmid UGCCIM library"
        /sex="Male"
        /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
        /note="Vector: pMD22my; Purified genomic DNA from M.
        musculus C57BL/6J (male) was obtained from the Jackson
        Laboratory Mouse DNA Resource
        (http://www.jax.org/resources/documents/dnares/). The DNA
        was hydrodynamically sheared by repeated passage through a
        0.005 inch orifice at constant velocity. The sheared DNA
        was blunt end-repaired with T4 DNA polymerase and T4
        polynucleotide kinase. Adaptor oligonucleotides were
        ligated to the blunt ends in high molar excess. The
        adaptor DNA was purified and size-selected for a 9.5 to
        10.5 kb range using preparative agarose gel
        electrophoresis. Vector DNA was prepared from a derivative
        of pMD22 (g114732114[9b]/AF123072.1), a copy-number
        inducible derivative of plasmid R1. The vector was ligated
        with adaptors complementary to the insert adaptors and
        purified. The sheared, adaptor mouse DNA was annealed to
        adaptor vector DNA, and transformed into
        chemically-competent E. coli XL10-Gold (Stratagene) cells
        and selected for ampicillin resistance."
BASE COUNT      133 a      113 c      141 g      93 t
ORIGIN
Query Match      6.7%; Score 82; DB 13; Length 480;
Best Local Similarity 58.8%; Pred. No. 3.9e-11;
Matches 161; Conservative 0; Mismatches 110; Indels 3; Gaps 1;

```



```

QY 810 tctctctgctcgtgtgttgcacatctgtgacccctccatctgacccgctctct 869
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 TTGACCTGTTGTGTATTCGGCATCTGCTGGGCTCCATTCATCTGACCTATCATGT 215

QY 870 tcaagctgtggaag--agtggacagaagtcctgtgctgtgtgttcaacctcatcag 926
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 214 GGAGCTGCTGTATGACACATCAACGGAAGGCTGCACCTGGCTACACAGTGTGTCCACA 155

QY 927 tgtatcagagtgctctcttctatctgaagtcgcggtcaacccctctctataaccc 986
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 154 TTGCTCTGGCAATCTTCTTATCTCGGCTCAGCAGCAACCCGGTGTCTTACAGCCCTCA 95

QY 987 tctctcggcgtctccggcgcccttcgcgaatgtgtctccctaccctcgaaatgtctcc 1046
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 94 TGTCTACTCGCTTCGGAAGACCTTCCTCGCAAGCCCTGGGCTTGGAACCCAGTGTCTGTC 35

QY 1047 atcccggcaccgcgcacaggaacctccagccca 1080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34 ATCGCCGCCAACCCATCATGCTTCCTCATGCGA 1

RESULT 14
LOCUS H11359 379 bp mRNA EST 26-JUN-1995
DEFINITION ym13d04.r1 Soares infant brain INIB Homo sapiens cDNA clone
IMAGE:47842 5' similar to contains L1 repetitive element ;, mRNA
sequence.
ACCESSION H11359
VERSION H11359.1 GI:876179
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 379)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
,M., Hulman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Ridkin,L., Rohlfing,T., Soares,M., Tan,F., Trevisakis,E., Waterston
,R., Williamson,A., Wohlmann,P. and Wilson,R.
The Washu-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 3363
High quality sequence stops: 321
Source: IMAGE Consortium, LNLN.
This clone is available royalty-free through LNLN; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert length: 3363 Std Error: 0.00
Seq primer: M13Rev
High quality sequence stop: 321.
Location/Qualifiers
1. 379
/organism="Homo sapiens"
/db_xref="GDB:420383"
/db_xref="taxon:9606"
/clone="IMAGE:47842"
/clone.lib="Soares Infant brain INIB"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lambda BA; Site_1: Not
I; Site_2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5];
AACGGAAGATTCGGCCCGCCAGCAATTTTCTTTTCTTTT 3');
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lambda BA vector."

```

```

BASE COUNT 124 a 63 c 74 g 115 t 3 others
ORIGIN

Query Match 6.3%; Score 77.8; DB 11; Length 379;
Best Local Similarity 79.8%; Pred. No. 4.7e-10;
Matches 91; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 722 aggcctgaagagatgaatcccttgagcgcaagaagtgtctgtaatactcagacc 781
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 266 AGNCTAAAGAAATCAATCTCTGAGCGAGATGAGGGAATATTCATAAGACCC 325

QY 782 tctagaagtcagtcacagaatgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 835
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 TCCAGAAATATCATCAACAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 379

RESULT 15
LOCUS CNS01XJ2 817 bp DNA GSS 12-MAY-2000
DEFINITION Tetradon nigroviridis genome survey sequence PUC-Or1 end of clone
202F23 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL171767.1 GI:7809824
VERSION AL171767
KEYWORDS GSS; genome survey sequence.
SOURCE Tetradon nigroviridis.
ORGANISM Tetradon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 817)
AUTHORS Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and
Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Unpublished
2 (bases 1 to 817)
REFERENCE Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fitzames,C., Wincker,P., Brotlier,P., Queller,F.,
Saurin,W. and Weissenbach,J.
Human gene number estimate provided by genome wide analysis using
Tetradon nigroviridis DNA sequence
Unpublished
3 (bases 1 to 817)
JOURNAL Direct Submission
REFERENCE Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
AUTHORS This sequence is a single read and was generated as part of a large
JOURNAL scale clone-and-sequencing project of the Tetradon nigroviridis
COMMENT genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetradon.
Location/Qualifiers
1. 817
/organism="Tetradon nigroviridis"
/db_xref="taxon:99883"
/clone="202F23"
/clone.lib="G"
/note="Genoscope sequence ID : C0AG202C12SP1-end :
PUC-Or1"
BASE COUNT 158 a 237 c 258 g 159 t 5 others
ORIGIN

Query Match 6.3%; Score 77.8; DB 13; Length 817;
Best Local Similarity 56.3%; Pred. No. 6e-10;
Matches 184; Conservative 1; Mismatches 138; Indels 4; Gaps 2;

QY 157 gctgctctccctggtggggtatggtgcacatctctgtgtgcatgtgtgtccgaca 216

```

```
Db 492 GCTGACGGGAATCGTGGGAAACGGGTGACATCAGGTGAGCCAGTGTGTAAGCAGAA 433
Oy 217 tcagacttgaagacaccccaactactatctctcagcttggcaagtctcagatctgct 276
Db 432 CCGCTACCTGCAGAACAGACGTCACCGACACATGATCAGCCCTGCTCCGACCTCCT 373
Oy 277 ggtcctgctcttgggagatgctcttggaaactca---cgaagatggcacatcacctt 333
Db 372 GGTGCTGCTCATCGCATGCGGTGAGCTCTACAGCGCCCATCTGTCCTTCACTTC 313
Oy 334 cctgtcgggacctgtggagtgtctactcaagacagccctcttggagactgtgtcttgc 393
Db 312 GCGTCCGCGCAACKCTCTGTAAATCTTCAAACTTCTGTTCGAGGC-GTGCAGTAGCG 254
Oy 394 ctccattctcaagtgtaacacaggttagcgtagagcgctatgtggcaatgtcaccctt 453
Db 253 CACCATTTGTGAACATAGCCACGCTTAGCTTCGAGCGCTACGTGGCCATCTGCCACCCCTT 194
Oy 454 ccgagccaagctggagagcaagcgagcg 480
Db 193 CCGCTACAAAGGCCCTGCGCGGAGCG 167
```

Search completed: April 22, 2002, 17:16:30
Job time: 10964 sec

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

Result	No.	Score	Query	Match	Length	DB	ID	Description
C	1	1200.4	97.5	134	10	AB041229	Rattus no	AB041229 Rattus no
	2	1186.4	96.4	1188	10	AF242875	Rattus no	AF242875 Rattus no
	3	799.6	65.0	1298	6	AF272363	Homo sapi	AF272363 Homo sapi
	4	799.6	65.0	1658	6	AX109234	Sequence	AX109234 Sequence
	5	799.6	65.0	1658	6	AX109236	Sequence	AX109236 Sequence
	6	796.4	64.7	1658	6	AX109238	Sequence	AX109238 Sequence
	7	796.4	64.7	1658	6	AX109240	Sequence	AX109240 Sequence
	8	790.8	64.2	1248	9	AB041228	Homo sapi	AB041228 Homo sapi
	9	790.8	64.2	1248	9	AF292402	Homo sapi	AF292402 Homo sapi
	10	788	64.0	1239	9	AF242874	Homo sapi	AF242874 Homo sapi
C	11	677.6	55.0	1594	6	AX109242	Sequence	AX109242 Sequence
	12	677.6	55.0	1594	6	AX109244	Sequence	AX109244 Sequence
	13	481.4	35.1	215088	2	AC008571	Homo sapi	AC008571 Homo sapi
	14	472.6	38.4	801	6	AX147772	Sequence	AX147772 Sequence
	15	469.8	38.2	729	6	AF273107	Sequence	AF273107 Sequence
	16	330.6	26.9	1318	9	AF272362	Homo sapi	AF272362 Homo sapi
	17	301.4	24.5	1209	10	AF248873	Rattus no	AF248873 Rattus no
	18	299.8	24.4	1239	10	AB038649	Rattus no	AB038649 Rattus no
	19	285.6	23.2	1218	10	AF044602	Mus muscu	AF044602 Mus muscu
	20	266	21.6	828	9	HSC0PC01	Homo sapi	AF044600 Homo sapi
C	21	266	21.6	168880	9	AC0017104	Homo sapi	AC0017104 Homo sapi
	22	151.2	12.3	759590	2	AC073449	Homo sapi	AC073449 Homo sapi
	23	141	11.5	1098	6	AX154586	Sequence	AX154586 Sequence
	24	133.2	10.8	1092	6	AR156357	Sequence	AR156357 Sequence
	25	133.2	10.8	1350	10	AB001982	Rattus no	AB001982 Rattus no
	26	132.8	10.8	1088	6	AR156353	Sequence	AR156353 Sequence
	27	132.8	10.8	1101	6	AX154584	Sequence	AX154584 Sequence
	28	132.8	10.8	1101	6	HS060179	Human growt	U60179 Human growt
	29	128	10.4	1063	6	AR156351	Sequence	AR156351 Sequence
	30	128	10.4	1101	4	SS060178	Sus scrofa	U60178 Sus scrofa
C	31	114.2	9.3	1676	5	AF082210	Spheroeold	AF082210 Spheroeold
	32	111.4	9.0	121652	34	AC008191	Drosophila	AC008191 Drosophila
	33	111.4	9.0	140838	2	AC018176	Drosophila	AC018176 Drosophila
	34	111.4	9.0	195212	3	AC007441	Drosophila	AC007441 Drosophila
	35	111.4	9.0	225374	2	AE003703	Drosophila	AE003703 Drosophila
	36	109.6	8.9	384	4	HSC0PC02	Homo sapi	AF044601 Homo sapi
	37	104.8	8.5	1161	6	AX154591	Sequence	AX154591 Sequence
	38	104.8	8.5	1239	6	AX154589	Sequence	AX154589 Sequence
	39	104.8	8.5	2040	9	AF034632	Homo sapi	AF034632 Homo sapi
	40	104.8	8.5	163284	9	AL137000	Human DNA	AL137000 Human DNA
C	41	104.8	8.5	341560	2	AL1596304	Homo sapi	AL1596304 Homo sapi
	42	103.4	8.4	836	6	AR156355	Sequence	AR156355 Sequence
	43	102.2	8.3	3129	6	AR156356	Sequence	AR156356 Sequence
	44	102.2	8.3	3129	10	RN099321	Rattus norw	U94321 Rattus norw
	45	100.2	8.1	870	9	HS060181	Human growt	U60181 Human growt

REFERENCE AUTHORS	TITLE	JOURNAL
1 (sites) Hosoya, M., Moriya, T., Kawabata, Y., Ohkubo, S., Fujii, R., Matsui, H., Shitani, Y., Fukusumi, S., Habata, Y., Hinuma, S., Onda, H., Nishimura, O. and Fujino, M.	Identification and Functional Characterization of a Novel Subtype of Neuremedin U Receptor	J. Biol. Chem. 275 (38), 29528-29532 (2000)

PUBMED 10887190
2 (bases 1 to 1314)
AUTHORS Fujii,R., Shintani,Y. and Hinuma,S.
TITLE Direct Submission
JOURNAL Submitted (03-APR-2000) to the DDBJ/EMBL/GenBank databases. Shuji
Hinuma, Takeda Chemical Industries, LTD, Discovery Research
Laboratories 1, Wadai 10, Tsukuba, Ibaraki 300-4293, Japan
(E-mail: Hinuma_Shuji@takeda.co.jp, Tel:81-298-64-5035,
Fax:81-298-64-5000)
FEATURES
source Location/Qualifiers
1..1314
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
127..1314
/gene="TGR-1"
127..1314
/gene="TGR-1"
/codon_start=1
/product="G-protein-coupled receptor TGR-1"
/protein_id="BAB13722.1"
/db_xref="GI:10257383"
/translation="MGKLENAIMHDPLKLYINTEEYLAHCGPKRDLSPVSVAY
ALFLVGVGNLNCMVIVRHOTLPPRYVYFSLAVSDILLLCMPEIYEMHNY
PLFPGVGCYFTALPEYVCFASISVTTVSERYVAIVHPRAKLESTRRALRLS
LWVSFVSFSLPNTSIHGIRFQHPGSSVPGSACTVTKPMYVNLIIQATSFLEYI
LPMTLISVLYLMDGLRKDESLKANKVAVNIHRSRKSVTMLFVLYVFAICMTPF
HYDRLEFSEVEMTESLAAVFNLIHVSVFVFLYSANVPIIYNLSRRFAFRNV
SPTCKMCHRHQPOGPAPAKIIFLRECHLMELTEDAGPQFPQSSIHNNLTMAPCAG
EVP"
BASE COUNT 269 a 397 c 305 g 343 t
ORIGIN
Query Match 97.5%; Score 1200.4; DB 10; Length 1314;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1204; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 gtctggaatttcaagctcaagtaattggaacttgaacttctctggaaccagatcc 60
DB 105 GTTGTGATTTTAAAGCTCAATATGGGAAACCTTGAAATGCTTCGATGCCAGATCC 164
QY 61 tctcagaagtaacttgaaacagacagaagaagtaacttggccaccctgtgtgaaccagag 120
DB 165 ACTCATGAAATGACTTGAACAGCAGAGGAGTACTTGGCCACCTGTGTGACCCAGCG 224
QY 121 cagtgacctaaccctccagtgctgtgagcctaagcctgacttctcctgtgtggagtaat 180
DB 225 CAGTGACCTATCCCTTCGCGGTGCTGTGGCTTATGCCCTGATCTTCTGTGGGGTAAAT 284
QY 181 gggcaactctctgtgtgtgacagtgatgtgcagacatacgaacttgaagacacacacaa 240
DB 285 GGGCAATCTTCTGTGTGATGTGATGTGCCAATCAACATTTGAAGACACCCACCA 344
QY 241 ctctactctctcagcttggcagctcagaatcgtcgtgtcctgtgtggagtgacct 300
DB 345 CTACTATCTCTTACGTGTGGCACTTCAGATCGCTGGTCTCTCTTGGGGATGCTCTT 404
QY 301 ggaatactgaagatgtgtgacaaattacaccttccctgttcggccctgtgagatcactt 360
DB 405 GGAATATCTGAGATGGGACAAATTACCTTCTCTGTGGCCCTGTGGAGTCACTCTT 464
QY 361 caagaagaccctctcgagaactgtgtgtcttgcctcactcagtgatcacaacggttag 420
DB 465 CAAGACAGCCCTCTTGAGAGCTGTGTGCTTGGCTCCTCATCTCAGTGTCAACAGGTTAG 524
QY 421 cgtagagcgtatgtgagcactgtgtcacccttccgagacgaagctgtgagacagcgag 480
DB 525 CGTAGAGCGCTATGTGGCCATTGTCCACCTTTCGAGCCAACTGTGAGAGCAGCGCGCG 584
QY 481 agcggccctcagaatcctcagctcagctcagctcctcgtgtgtgttcttcttgcacaa 540
DB 585 ACGGGCCCTCAGATCCTCAGCTAGCTGTGAGACTTCTCTGTGTCTTCTTGTGCCCAA 644

QY 541 taccagatccatgagatcagatcagcttccagcacttcccaacgggtccctcgtacctgctc 600
DB 645 TACCAGATCCATGATGATCAAGTTCCAGCACTTTCACCAAGGGCTCCGTACTGACTG 704
QY 601 agcgaactgtcagatcagatcagacacacacacacacacacacacacacacacacacac 660
DB 705 AGCCACTGACAGATCAACCAACCACTGTGGGTATTAACCTTATCATCCCACTACAG 764
QY 661 ctctctcttctacatccctcccaatgacccctcagcagcgtctcctactactcctcagggct 720
DB 765 CTCTCTCTTCTACATCTCTCCCATGACCCCTCATCAGAGTCTCTACTACTCATTTGGGGCT 824
QY 721 cagagctgaagaagatgaatcccttgaaggcgaacaaatgtgctgtgaatatcagagacc 780
DB 825 CAGGCTGAAGAGATGAATCCCTTGAGCGCAACAAAGTGTGATATTCACAGAGACC 884
QY 781 ctctagaagatcagatcagac 840
DB 885 CTCTAGAAAGTCAGTCAACCAAGATGCTGTGTGTGCTCTGTGTGTGCTGTGTGCTGTG 944
QY 841 gaccctctcagatgtgagccgctcctctcagcttgtgaaagatgagacagatccct 900
DB 945 GACCCCTTCATGATGTGACCGGCTCTTCTTACCTTGTGTGAAGAGTGACAGATCCCT 1004
QY 901 ggtctgtgtgtcaacatcagatcagatgagatgagatgagatgagatgagatgagatgag 960
DB 1005 GGTCTGTGTGTCAACCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 1064
QY 961 ggtcaacccatatactatactcctcgtctgtgagccttccggggcgcttctgaaatgt 1020
DB 1065 GGTCAACCCCAT 1124
QY 1021 tgtctccctcactgcaaatgtgtgcaatcccgagatcagcagagacccagaccca 1080
DB 1125 TGCTCTCCCTCACTGCAAAATGATGCTCCCTCCGATCAGCCACAGGACTCTCAGCCCA 1184
QY 1081 gaagatcattcttgcagagaatgtcacctcgtgtgagatgtgacagaagatgacagcccca 1140
DB 1185 GAAGATCATCTTCTGTGACGAATGTCACTCATGATGATGATGATGATGATGATGATGATG 1244
QY 1141 gttccctgtgtcagatcagatcagacacacacacacacacacacacacacacacacac 1200
DB 1245 GTTCCCTGTGTCACTCATGATGATGATGATGATGATGATGATGATGATGATGATG 1304
QY 1201 ggtacacata 1210
DB 1305 GTTACCATGA 1314
RESULT 2
AF242875 1188 bp mRNA ROD 13-JUL-2000
LOCUS Rattus norvegicus neuromedin U receptor 2 (NMU2R) mRNA, complete
DEFINITION
ACCESSION AF242875
VERSION AF242875.1 GI:9082157.
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1188).
AUTHORS Howard,A.D., Wang,R., Pong,S.S., Mellin,T.N., Strack,A., Guan,X.M.,
Zeng,Z., Williams,D.L., Jr., Feiguer,S.D., Nunes,C.N., Murphy,B.,
Stalir,J.N., Yu,H., Jiang,Q., Clements,M.K., Tan,C.P., McKee,K.K.,
Hreniuk,D.L., McDonald,T.P., Lynch,K.R., Evans,J.F., Austin,C.P.,
Caskey,C.T., Van der Ploeg,L.H. and Liu,Q.
TITLE Identification of receptors for neuromedin U and its role in
feeding
JOURNAL Nature 406 (6791), 70-74 (2000)
MEDLINE 20351041
REFERENCE 2 (bases 1 to 1188)

AUTHORS Liu, Q., McDonald, T. P., Wang, R., Jiang, Q. and Howard, A. D.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) Pharmacology, Merck Research Labs, West
Point, PA 19486, USA
FEATURES Location/Qualifiers
source 1. .1188
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
gene 1. .1188
/gene="NMUR2"
CDS 1. .1188
/gene="NMUR2"
/note="FM-4"
/codon_start=1
/product="neuromedin U receptor 2"
/protein_id="AA82736.1"
/db_xref="GI:9082158"
translation="MGKLENSAWIHDPMLKYNSTEXLAHLCGPKRSDSLPVSVA
ALIFVGMGLVCMVIRHQTPTNRYLFSLAVSLDVLGLMPLLEYEMHNY
PLFPGVGYKTALEFVVCFASTLSVTVEVVAIVHFRKLESTRRRLRILS
LYMSFVSFSLPNTSINGIKRQHPNNSVSGATCTYTRKMYNLIVQTSFLFI
LPMFLISVLTLMGLRKDESLANKVAIVIHPSRSVTKMLFVLVAICWTFP
HVDLFEFSVEEMTESLAAYFNLIHVGVEFYLSAVNPITVNLISRFRAARMV
SPCKMCHPRHPOGPPAQKIIFLTECHVELTBDAGFQFOSSIHNTNLTPACAG
EVP"
BASE COUNT 243 a 361 c 278 g 306 t
ORIGIN
Query Match 96.4%; Score 1186.4; DB 10; Length 1188;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1187; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 23 atgggaaacttgaaatgctctcgatcagatcctctcatgaagtacttgaaacagc 82
Db 1 ATGGGAAACTTGAATAAGCTTCTCGATCCAGATCCACTCATGAACTTGAACAC 60
QY 83 acagaggagtacttggccaactctgtgagacccaagcgagacctctcccttcggg 142
Db 61 ACAGAGGAGTACTTGGCCCACTGTGTGGACCAAGCCAGTACCTATCCCTCCGGG 120
QY 143 tctgtgacctatgagcttctctctgtgtggtgaaatgagcaactctctgtgcatg 202
Db 121 TCTGTGGCTATATGGCTATCTCTGTGGGGGTAATGCGAATCTTGTGTGCAAG 180
QY 203 gtgattgtcgaacatcagaactttgaagaacaccacaactactatctcttcagcttg 262
Db 181 GTGATTGTCCGACATCAGACTTTGAAGACACCACCACTACTATCTTTCAGCTTGGCA 240
QY 263 gtctcagatctgtgtctgtctctgtgtggtggtcctcttgaaatctagaagtgtggcgc 322
Db 241 GTCTCAGATCTGTCTGTCTCTTGTGGGGATGCTCTGGAATCTACAGATGTGGCAGC 300
QY 323 aattacccttctctgtcgggctgtgagatgtaacttcaagaacagccctctcgaagct 382
Db 301 AATTACCTTTCCTTGTGGGCTGTGGGATCTACTTCAACAGACCCCTCTTCAGACT 360
QY 383 gtgtgcttctgacctcattctcagtgatccacaggttagcgttagagcgctatgttgccatt 442
Db 361 GTGTCTTGTGCTCATCTCTCAGTCTCACACAGGTTAGGCTAGCGGTATGTGGCATT 420
QY 443 gtccacccttctcggagccaagcttgagagacagcgagcgagccctcagaatctctcagc 502
Db 421 GTCCACCTTTCGAGCAAGCTGAGAGCAAGCGCGCCCTCAGGATCTTCACAC 480
QY 503 ctgagcttggaagctctcgtgtgtcttctcttgcaccaatcagatcagatgagcaag 562
Db 481 CTAGCTGTGAGAGTCTCTGTGTCTTCTTTCGCCAATACAGACATCATGTGCAATCAAG 540
QY 563 ttcacgaacttcccaagcggtcctcgttaccgtgctcagccacctgacagctcaccaaa 622
Db 541 TTCAGACACTTCCCAAGGCGTCTCGTACCTGCTCAGCCACCTGCACAGTCAACAAA 600

QY 623 cccatgtgggtgtataacttgatcatccaagctaccagcttctcttactacatctccca 682
Db 601 CCCATGTGGGTATATTAATTGATCATCCAGCTACCAAGCTTCTCTTCTACTATCTCCCA 660
QY 683 atgacctatcagcttctctactactacatcagtgaggctcgaagctgaagaagatgaatcc 742
Db 661 ATGACCTCATCAGCGTCTCTACTACCTCATCGGGGCTCAGGCTGAAGAGATGAATACC 720
QY 743 ctgtagcggaacaaagtggctgtgaaatattcaagacctctagaagaagtcagcaaga 802
Db 721 CTYTAGCGGCAACAAAGTGTGTAATATTCACACACCTCTAGAAAGACATCCACCAAG 780
QY 803 atgctgtgtgtctgtgctctgtgtgtgacatctcgtgaagcccttcacatgtgagcg 862
Db 781 ATGCTGTTTGTCTGTGTCTCTGTGTGTGCAATCTCTGAGACCCCTTCATGTGACCG 840
QY 863 ctctctctcagcttctgtgaagaagtgcagaagctcctcgtcgtgtgttcaacctcacc 922
Db 841 CTCTTCTTACGCTTGTGTGAAGAGTGCAGAGAGTCCCTGCTGCTGTTCAACTCATAC 900
QY 923 catgtgtatcagatgtgtctcttattatctgagctcgggtgtaacccatctataac 982
Db 901 CATGTGTATCAGGTGTCTTCTTATCTGACCTCCGCGTCAACCCCATTTATATATAC 960
QY 983 ctccctgtcggcgcttccggcgcgcttctgaaatgtgtctccctcaccctgcaaatgg 1042
Db 961 CTCTGTCTCGCGCTTCCGGCGGCTTTCGAATGTGTCTCCCTTACCTGTCAAAATGG 1020
QY 1043 tgcacatcccgcatcgcgcacagaaggacctccagccagaagaatcatctcttcagaga 1102
Db 1021 TGCATTCGCCGATCGGCACAGGAGACTCCAGCCACAAATATCATCTTGTGACAGAA 1080
QY 1103 tgtcaactctgtgagatctgcagaagatgcaggccccagcttccctggttgcattccatc 1162
Db 1081 TGTCACTCTGTGACTGCAGAGATGCAGGCCCCCACTTCCCTGCTCAGTATCATC 1140
QY 1163 caacaacacaaacttaccacagccccctgtgcaggagaagatccataa 1210
Db 1141 CACAACACCACTTACCAAGGCCCCCTGTGCAGAGAGATCACTAA 1188
RESULT 3
AF272363
LOCUS AF272363 1299 bp mRNA PRI 23-OCT-2000
DEFINITION Homo sapiens neuromedin U receptor 2 (NMUR2) mRNA, complete cds.
ACCESSION AF272363
VERSION AF272363.1 GI:10946202
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1298)
Raddatz, R., Wilson, A.E., Artymyshyn, R., Bonini, J.A., Borowsky, B.,
Boreju, L.W., Zhou, S., Kouranova, E.V., Ngorny, R., Guayra, M.S.,
Dal, M., Lerman, G.S., Vayse, P.J., Branchek, T.A., Gerald, C.,
Porray, C. and Adam, N.
Identification and Characterization of Two Neuromedin U Receptors
Differentially Expressed in Peripheral Tissues and the Central
Nervous System
JOURNAL J. Biol. Chem. 275 (42), 32452-32459 (2000)
PUBMED 10899166
REFERENCE 2 (bases 1 to 1298)
Bonini, J.A., Raddatz, R., Wilson, A. and Borowsky, B.
Direct Submission
TITLE Submitted (25-May-2000) Target Discovery and Assessment, Synaptic
Pharmaceutical Corporation, 215 College Road, Paramus, NJ 07652,
USA
FEATURES Location/Qualifiers
source 1. .1298
/organism="Homo sapiens"
/db_xref="taxon:9606"

[illegible]

LOCUS	1248	bp	mRNA	PRI	21-SEP-2000
DEFINITION	AB041228 Homo sapiens mRNA for G protein-coupled receptor TGR-1, complete cds.				
ACCESSION	AB041228				
VERSION	AB041228.1 GI:10257380				
KEYWORDS	G protein-coupled receptor TGR-1.				
SOURCE	Homo sapiens CDNA to mRNA.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (sites) Hosoya,M., Moriya,T., Kawamata,Y., Ohkubo,S., Fujii,R., Matsui,H., Shintani,Y., Fukusumi,S., Habeta,Y., Hinuma,S., Onda,H., Nishimura,O. and Fujino,M.				
TITLE	Identification and Functional Characterization of a Novel Subtype of Neuropeptide Y Receptor				
JOURNAL	J. Biol. Chem. 275 (38), 29528-29532 (2000)				
REFERENCE	2 (bases 1 to 1248)				
PUBMED	10887190				
AUTHORS	Shintani,Y., Moriya,T., Ohkubo,S. and Matsui,H.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-APR-2000) to the DDBJ/EMBL/GenBank databases. Yasushi Shintani, Takeda Chemical Industries, LTD, Discovery Research Laboratories 1: 10 Madai, Tsukuba, Ibaraki 300-4293, Japan (E-mail:Shintani.Yasushi@takeda.co.jp, Tel:81-298-64-5011, Fax:81-298-64-5000)				
FEATURES	Location/Qualifiers				
source	1..1248				
gene	/organism="Homo sapiens"				
cds	/db_xref="taxon:9606"				
	10..1248				
	/gene="TGR-1"				
	10..1248				
	/gene="TGR-1"				
	/codon_start=1				
	/product="G protein-coupled receptor TGR-1"				
	/protein_id="BAB13721.1"				
	/db_xref="GI:10257381"				
	/translation="MEKLVNNSWYQOKLEDPFKQHLNSTEELYLAFICGRSRSHFLP				

[illegible]

[illegible][illegible]

QY	1022	gctcccccactgcaaaagtgccatcccccgcacgcgcgaagagactccagccag	1081
Db	1024	AACTCTCTTCCACAAAGAGTGGCACTCCACACAGACCACTGGCACCCTCTG	1083
QY	1082	aagatcattctctgacagaatgtcacctcgttgagactgacagaagatgacagccacag	1141
Db	1084	CGGAAACCTTTCTCTGACAGAAATGCCACTTGTGTGAGCTGACCGAAGATATATGATCCCA	1143
QY	1142	ttccctgctcagtcacatccatccacaacacacttaccacagccctctgtgagagag	1201
Db	1144	TTCCCATGTGACATTCATGCACAACTCTCACCCTCCCAAGCCCTCTCTAGTGAACAG	1203
QY	1202	gtaccataaa 1211	
Db	1204	ATGTCAAGAA 1213	
RESULT	10		
LOCUS	AF242874	1239 bp	PRI 13-JUL-2000
DEFINITION	Homo sapiens neuromedin U receptor 2 (NMU2R) mRNA, complete cds.		
ACCESSION	AF242874		
VERSION	AF242874.1	GI:9082155	
KEYWORDS	human.		
SOURCE	ORGANISM	Homo sapiens	
REFERENCE	1 (bases 1 to 1239)	Howard, A.D., Wang, R., Pong, S.S., Mellin, T.N., Strack, A., Guan, X.M., Zeng, Z., Williams, D.L., Jr., Feigheimer, S.D., Nunes, C.N., Murphy, B., Stalar, J.N., Yu, H., Jlang, Q., Clements, M.K., Tan, C.P., McKee, K.K., Hreniuk, D.L., McDonald, T.P., Lynch, K.R., Evans, J.F., Austin, C.P., Caskey, C.T., Van der Ploeg, L.H. and Liu, Q.	
AUTHORS	Identification of receptors for neuromedin U and its role in feeding		
TITLE	Nature 406 (6791), 70-74 (2000)		
JOURNAL	20351041		
MEDLINE	2 (bases 1 to 1239)		
REFERENCE	Liu, Q., McDonald, T.P., Wang, R., Jlang, Q. and Howard, A.D.		
AUTHORS	Direct Submission		
JOURNAL	Submitted (09-MAR-2000) Pharmacology, Merck Research Labs, West Point, PA 19486, USA		
FEATURES	Location/Qualifiers		
source	1..1239		
gene	/organism="Homo sapiens"		
CDS	/db_xref="taxon:9606"		
	/chromosome="5"		
	1..1239		
	/gene="NMU2R"		
	1..1239		
	/note="FM-4"		
	/codon_start=1		
	/product="neuromedin U receptor 2"		
	/protein_id="AAF8275.1"		
	/translation="MEKLONASMIYQOKLEDPFKHLNSTEELAFILCGPRSHFPLVSVVPIEVGVGIVGNVAVLCVILIOHAKKPTNYLSELAISDLIVILIGMPLFVEWMKRNPLFLGPGVGCYFKRALFETVCPASLITSTTVSVERVAIILHFRKLLQSTRRAKRLTIGVGFVSLPNTSIHGIRKFEHPNSLVPGSAWTCVIRKMWLYNLTIOVTSFLFYLPMVIVLSVLYLMAIRLAKKSLSEADGNANIQPRCKRSVKMKLFVLVFAICWAPHLIRLFEFVEWESLSAAVNLVAVGVFYLSSAVNPITLYNLSRFPAAFONVTSFHKWHSOHDPOLPPAQRNIIFLECHFVELTIDIQPFPCSSMNSHLPALSSPSMRNTNQSHPKNT"		
BASE COUNT	267 a 392 c 265 g 315 t		
ORIGIN			
Query Match	64.0%; Score 788; DB 9; Length 1239;		
Best Local Similarity	79.2%; Pred. No. 1.3e-207;		
Matches	954; Conservative 0; Mismatches 235; Indels 15; Gaps 1;		


```

Consensus quality: 204117 bases at least Q40
Consensus quality: 210928 bases at least Q30
Consensus quality: 212668 bases at least Q20
Estimated insert size: 222000; pulse-field gel estimation
Estimated insert size: 247880; sum-of-contigs estimation
Quality coverage: 6.63 in Q20 bases; pulse-field gel estimation
Quality coverage: 6.89 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 20385: contig of 20385 bp in length
* 20386 20485: gap of unknown length
* 20486 35143: contig of 14658 bp in length
* 35144 35243: gap of unknown length
* 35244 84715: contig of 49471 bp in length
* 84715 84814: gap of unknown length
* 84815 215088: contig of 130274 bp in length.
Location/Qualifiers

```

```

source
1. .215088
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="C1C-550M4"
/clone_1fb="Caltech human BAC library C"
41474 c 42096 g 64807 t
BASE COUNT 66410 a
ORIGIN

```

Query Match	39.1%	Score 481.4	DB 2	Length 215088
Best Local Similarity	79.6%	Pred. No. 1.9e-122		
Matches 388	Conservative	0	Mismatches 136	Indels 15; Gaps 1

[illegible]

OY	530	tccttggcccaataccagatattcaatggatcaagtctcagcaactttcccaacggtctctc	569
Db	105030	ttccctggcccaacaccagatattcaatggatcaagtctcagcaactttcccaacggtctctc	1049711
OY	590	gtacctgtgtcagcaaccctgtcacagatcaaccacaatgtgtgtgtataactgcatc	649
Db	104970	gtcccaagtttggggcactctgtacggtctatcagaccatgttgatcttcaatttcattc	10491111
OY	650	caagctacagacttctctctctacatccccaatgaaccatcaacgctctctactac	709
Db	104910	caggtacacactac	1048511
OY	710	ctcaatgggtctcagcttga	728
Db	104850	ctcatggcactctcagcttga	104832

RESULT	14		
AX147772/c			
LOCUS	AX147772	801 bp	DNA
DEFINITION	Sequence 17 from Patent WO0136473.		PAT
ACCESSION	AX147772		
VERSION	AX147772.1	GI:14346807	
KEYWORDS			

ORGANISM

REFERENCE

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi; Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo. I (bases 1 to 801)

Vogeli, G., Wood, L.S., Parodi, L.A., Hiesch, R.R., Lind, P., Vogeli, G., Wood, L.S., Kayes, P.S., Bannigan, C.M., Ruff, V., Slomcom, J., Schellin, K.A., Kayes, P.S., Bannigan, C.M., Ruff, V.,

AUTHORS

TITLE	Novel g protein-coupled receptors
JOURNAL	Patent: WO 0136473-A 17 25-MAY-2001
FEATURES	PHARMACIA & UPJOHN COMPANY (US)
	Location/Qualifiers

BASE COUNT	200	a	187	c	261	g	15
ORIGIN	/organism="Homo sapiens" /db_xref="taxon:9606"						

Query Match	38.4%	Score 472.6	DB 6	Length 801
Best Local Similarity	79.5%	Pred. No. 4.2e120		
Matches 578	Conservative 0	Mismatches 134	Indels 15	Gaps 1

Oy	17	tcagaaatcgggaaaacttgaaaatgctctccgagtcac-----gatcct	61
Db	801	TCAGGATGGAAAACTTCACAAATGCTCTTCGATCTTACACGACGAANACTGAAAGATCCA	742
Oy	62	ctcactgaagtacttgaacagcacagagagtagcttggccaccctgtgttgaccacaagcgc	121
Db	741	TTCCGAAACACCTGAMCAGCACCGAGAGATATGCGCTTCCTCGGACCTCGGCGC	688
Oy	122	agtagccatacccttcgcggtctcgttgagccatcgcctgactctccgcgtggtgggtatg	181
Db	661	AGCCACTTCTCTCCGCCGTCCTGTGTGTATCTGCAATTTTGTGGTGGGGGTATT	622
Oy	182	ggcaactctctcgtgtgtgcatgtgatgttcggacaatcagaacttgaagaacaccacaac	241
Db	621	GGCAATGTCTGTGTGTGCTGTGTATCTGACGACACGAGGTATGAAAGACGCCACCAAC	562
Oy	242	tactatctcttcacgtctggcagaatctcagatctcgtctgtctctctctgggagatgcctctg	301
Db	561	TACTACCTCTTCACGCTGTGGCTGTCTGACCTCTCTGTGCTGCTCTTGGAAATGCCCCTG	502
Oy	302	gaaatctcagagatgtggacaatatccacttctcgtctcgtcggccctgtgtggagtactctc	361
Db	501	GAGGCTATGAGATGTGGCGCACTACCCCTTCTTGTGTGGGCCCGGTGGCTGTACTTC	442
Oy	362	aagacagccctcttcagagatgtgtgtctgtccatcttcagtgctacacacggttacc	421

```

Db 441 AACAGCGCCCTTTTGAGACCGTGTGCTTCGCCCTCCATCCATGACATACACACCGTACG 382
QY 422 gttagagcgtatgtggtgcatgtgtcacccttccagagcaagtgtagagcagcgagga 481
Db 381 GTTGAGACGGCTACGTGGCCATCTTACACCCGTTCCGGCCCAACTGACAGACACCGCGC 322
QY 482 cgggacctcaagatccctcagcagctgtgtagcttctgtgtgtcttcttcttcccaat 541
Db 321 CGGGCCCTGAGATCCTCGGCAATCGTGTGGGGCTTCGCCGCTCTTCCTTCCTCCGCCAAC 262
QY 542 accaagatcagatgtgcatcaagttccagcagacttcccaaggggtccctccgtactgtctca 601
Db 261 ACCAGATCCATGGCATCAAGTTCCACTACTTCCCAATGGGTCCCTGGTCCAGGTTCCG 202
QY 602 gcaacctgcaagtgcaaccaaaccaatgtgtgtgtatgaacttgatccaagctaccagc 661
Db 201 GCCACCTGTACGGTTCATCAAGCCCATGTGATCTACATTTCTCATCCAGGTCACCTCC 142
QY 662 ttccctctctacatcctcccaatgaccctcaccagcgtccctcactacatctatgtggctc 721
Db 141 TTCTATTCTACTCTCTCCCATGACTGTCATCATGTGTCTCTACTACTACTGACGACCTC 82
QY 722 aggtctga 728
Db 81 AGAGTGA 75

```

RESULT 15

```

LOCUS AX139107 729 bp DNA PAT 30-MAY-2001
DEFINITION Sequence 1 from Patent EP1090990.
ACCESSION AX139107
VERSION AX139107.1 GI:14274786
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 729)
AUTHORS Harland, L. P.
TITLE Human g-protein-coupled receptor
JOURNAL Patent: EP 1090990-A 1 11-APR-2001;
Pfizer Limited (GB); PFIZER INC. (US)
FEATURES
source 1..729
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

```

BASE COUNT 135 a 245 c 167 g 182 t
ORIGIN

```

Query Match 38.2%; Score 469.8; DB 6; Length 729;
 Best Local Similarity 79.6%; Pred. No. 2.5e-119;
 Matches 574; Conservative 0; Mismatches 132; Indels 15; Gaps 1;

```

QY 23 atgggaacttgaanaatgtctcctggtatccac-----gattccttcattg 67
Db 1 ATGGAAAAATTCAGATGCTTCTCTGATCTTACAGAGAACTAGAAAGATCCATTCCAG 60
QY 68 aagtaacttgaacacagagagagtaacttggccaccctgtgtgacccaagcgcaatgac 127
Db 61 AAACACCTGAACAGACAGCGAGAGATATCTGCTTCTCTGCGACCTCGCGCAGCGAC 120
QY 128 ctatcccttcgggtgtctgtgtgacctatgctatctcctgtgtgtgggttaatgtgcaat 187
Db 121 TTCTTCTCTCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 188 ctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 247
Db 181 GTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
QY 248 ctcttcagcttgcagttctcagatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 307

```

```

Db 241 CTCTTAGCGCTGGCGGCTCTCTGTGACCTCTGTGCTCTGCTCTCTGGAATGCCCTGAGAGTTC 300
QY 308 taagagatgtgtgacaataataacccttccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 367
Db 301 TATGAGATGTGGGCAACTACTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
QY 368 gcccttcgagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 427
Db 361 GCCCTTTTGTGAGACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
QY 428 cgtatgtgtgcatgtgtcacccttccgagaccagctgtgtagagcaagcgagcgagcc 487
Db 421 CGCTAGTGTGCGCATCTCTACACCGGTTCCGGCCAAATGCAAGACCCGCGCCGCGGCC 480
QY 488 cttagatcctcagcctagctgtgtagcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 547
Db 481 CTGAGATCTCTGGCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
QY 548 atccatgtgcatcaagttccagcacttcccaacgggttccctcgttaccgtgtcagcacc 607
Db 541 ATCCATGTGCAATCAAGTTCCACTACTTCCCAATGGGTCCCTGGTCCAGGTTCCGCGCAC 600
QY 608 tgcacagttcccaaacccatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 667
Db 601 TGTACGTCATCAAGCCCATGTGATCTACAAATTTATCATATCCAGGTCACGCTTCTTCC 660
QY 668 ttctacatcctcccaatgaccctcaccagcgttccctcactacatctatgtgggtcagcgctg 727
Db 661 TTCTACTCTCTCCCATGACTGTCATCATGTGTCTCTACTACTCATGGCAGTCAAGAGTG 720
QY 728 a 728
Db 721 A 721

```

Search completed: April 22, 2002, 18:22:55
 Job time: 11:34 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:13:42 ; Search time 43.26 Seconds
(without alignments)
1403.214 Million cell updates/sec

Title: US-09-609-146-4

Perfect score: 2185

Sequence: 1 MSQMEKLQNSMYIOOKLED.....ALSSQMSRTNYGSPFNKR 415

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL_17:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2185	100.0	415	4	09GZ04
2	2170	99.3	412	4	09NRA6
3	1650.5	75.5	395	11	09JTB1
4	1643.5	75.2	395	11	09ES04
5	942	43.1	403	4	043664
6	942	43.1	426	4	09HBB9
7	885.5	40.5	402	11	09JTB2
8	880.5	40.3	412	11	09JTB5
9	825.5	37.8	405	11	05S040
10	552.5	25.3	428	5	09VEN4
11	515	23.6	595	5	09VFW6
12	497.5	22.8	660	5	09VFW5
13	480	22.0	378	5	018701
14	470.5	21.5	418	5	017239
15	469	21.5	363	13	093413
16	444.5	20.3	374	13	093412
17	424	19.4	397	13	09DDR1
18	390	17.8	401	13	09DDR0
19	361.5	17.5	404	13	09DFB0

20	378	17.3	559	13	093414	093414 spherooides
21	376	17.2	513	13	09DFA9	09df9 catostomus
22	371.5	17.0	377	13	098U14	098u14 brachydantio
23	361	16.5	383	13	042324	042324 catostomus
24	360	16.5	438	11	09R0D1	09r0d1 mus musculu
25	356.5	16.3	352	11	088820	088820 rattus norv
26	356.5	16.3	352	11	09R297	09r297 rattus norv
27	356.5	16.3	352	11	09QW03	09qw03 rattus norv
28	356.5	16.3	452	11	09JKN0	09jkn0 mus musculu
29	355.5	16.3	393	11	09R1M0	09r1m0 mus musculu
30	355.5	16.3	401	11	09R1L9	09r1l9 mus musculu
31	355.5	16.3	444	11	09JTY1	09jty1 mus musculu
32	355	16.2	454	4	09H573	09h573 homo sapien
33	354	16.2	400	6	09MYW9	09myw9 macaca mula
34	351	16.1	464	5	09GQ54	09gq54 aedes aegypt
35	348	15.9	504	5	09VGX8	09vgx8 drosophila
36	347	15.9	394	5	09U721	09u721 drosophila
37	346.5	15.9	367	13	09PVG0	09pvg0 carassius a
38	346.5	15.9	367	13	09PVP9	09pvp9 carassius a
39	346	15.8	384	13	098UH1	098uh1 brachydantio
40	346	15.8	411	13	09W6T3	09w6t3 gallus gall
41	345	15.8	388	4	09UTY1	09uty1 homo sapien
42	341.5	15.6	385	11	09JK40	09jk40 mus musculu
43	339	15.5	398	4	09UDE7	09ude7 homo sapien
44	336	15.4	429	4	013729	013729 homo sapien
45	336	15.4	455	4	060451	060451 homo sapien

ALIGNMENTS

RESULT 1
ID 09GZ04 PRELIMINARY: PRT: 415 AA.
AC 09GZ04:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE NEUROMEDIN U RECEPTOR 2 (NEUROMEDIN U RECEPTOR-TYPE 2).
GN NMUR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20490668; PubMed-10899166;
RA Raddatz R., Wilson A.E., Artymyshyn R., Bonini J.A., Borowsky B.,
RA Boteju L.W., Zhou S., Kouranova E.V., Ngorny R., Guevarra M.S.,
RA Dai M., Lerman G.S., Vayse P.J., Branchek T.A., Gerald C., Forray C.,
RA Adham N.;
RT "Identification and Characterization of Two Neuromedin U Receptors
RT Differentially Expressed in Peripheral Tissues and the Central Nervous
RT System.";
RL J. Biol. Chem. 275:32452-32459(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed-11010960;
RA Shan L., Qiao X., Crona J.H., Behan J., Wang S., Laz T., Bayne M.,
RA Gustafson E.L., Monsma F.J. Jr., Hedrick J.A.;
RT "Identification of a Novel Neuromedin U Receptor Subtype Expressed in
RT the Central Nervous System.";
RL J. Biol. Chem. 275:39482-39486(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Pang L., Wang S., Laz T., Hedrick J.A.;
RT Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF272363; AAG24794.1; -
DR EMBL: AF292402; AAG03064.1; -
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSN.
DR PROSITE: PS00237; G_PROTEIN_RECPT_FL_1; UNKNOWN_1.

DR PROSITE: PS50262; G_PROTEIN_RECCEP_FL_2; 1.
 KM RECEPTOR.
 SQ SEQUENCE 415 AA: 47725 MW: C2BACD84B313390F CRC64:

Query Match 100.0%; Score 2185; DB 4; Length 415;
 Best Local Similarity 100.0%; Pred. No. 2.5e-154;
 Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSGMEKLNASWYQOKLEDPEFKHLNSTEETVLAFLCGPGRSHFFLPVSVVYPIFVVG 60
 DB 1 MSGMEKLNASWYQOKLEDPEFKHLNSTEETVLAFLCGPGRSHFFLPVSVVYPIFVVG 60
 OY 61 IGVNLVCLVILQHQAMKPTPTNYLFLSLAVSDLLVLLGMPLEVEYEMRNYPFLFGVGCY 120
 DB 61 IGVNLVCLVILQHQAMKPTPTNYLFLSLAVSDLLVLLGMPLEVEYEMRNYPFLFGVGCY 120
 OY 121 FRTALFETVCFASILSTTVSVERYVAILHPFRAKLQSTRRRRLRLGLGWGFSVLSPLP 180
 DB 121 FRTALFETVCFASILSTTVSVERYVAILHPFRAKLQSTRRRRLRLGLGWGFSVLSPLP 180
 OY 181 NTSIHGKFPNGSLVPGSATCTVIKPMWYINFTIOVTSFLFYLLPMTVISLVLYLMA 240
 DB 181 NTSIHGKFPNGSLVPGSATCTVIKPMWYINFTIOVTSFLFYLLPMTVISLVLYLMA 240
 OY 241 LRLKDKSLAEDEGNANIORPCRSYKMKLFVLVLFALICWAPFHIDRLFFSFVEEMSES 300
 DB 241 LRLKDKSLAEDEGNANIORPCRSYKMKLFVLVLFALICWAPFHIDRLFFSFVEEMSES 300
 OY 301 LAAVNLVHVSGVFYLSAVNPITYNLLSRFOAFQNVISSEFKQWHSQDLPOLPPA 360
 DB 301 LAAVNLVHVSGVFYLSAVNPITYNLLSRFOAFQNVISSEFKQWHSQDLPOLPPA 360
 OY 361 QNIFTECHFEVLETDIGPOFPCOSSMHNHSLPTALSSQMSRTYQSFHFNKT 415
 DB 361 QNIFTECHFEVLETDIGPOFPCOSSMHNHSLPTALSSQMSRTYQSFHFNKT 415

RESULT 2

O9NRA6 PRELIMINARY: PRT: 412 AA.

AC 09NRA6: 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE NEUROMEDIN U RECEPTOR 2 (G PROTEIN-COUPLED RECEPTOR TGR-1).
 GN NMUR2R OR TGR-1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20351041; PubMed=10894543;
 RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
 RA Zeng Z., Williams D.L., Feigheimer S.D., Nunes C.N., Murphy B.,
 RA Stair J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McKee K.K.,
 RA Hrenluk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austin C.P.,
 RA Caskey T., van der Ploeg L.H.T., Liu Q.;
 RT "Identification of receptors for neuromedin U and its role in
 feeding.";
 RL Nature 406:70-74(2000).
 RP [2]
 RP SEQUENCE FROM N.A.
 RX PubMed=10887190;
 RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
 RA Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
 RA Fujino M.;
 RT "Identification and Functional Characterization of a Novel Subtype of
 Neuromedin U Receptor.";
 RL J. Biol. Chem. 275:29528-29532(2000).
 DR EMBL: AF242874; AAF82755.1; -;
 DR EMBL: AB041228; BAB13721.1; -;

DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCR_RHODPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECCEP_FL_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECCEP_FL_2; 1.
 KM RECEPTOR.
 SQ SEQUENCE 412 AA: 47450 MW: 333D6D1A4841B63E CRC64:

Query Match 99.3%; Score 2170; DB 4; Length 412;
 Best Local Similarity 100.0%; Pred. No. 3.2e-153;
 Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 MEKLNASWYQOKLEDPEFKHLNSTEETVLAFLCGPGRSHFFLPVSVVYPIFVVGIGN 63
 DB 1 MEKLNASWYQOKLEDPEFKHLNSTEETVLAFLCGPGRSHFFLPVSVVYPIFVVGIGN 60
 OY 64 VLVCVILQHQAMKPTPTNYLFLSLAVSDLLVLLGMPLEVEYEMRNYPFLFGVGCYFKT 123
 DB 61 VLVCVILQHQAMKPTPTNYLFLSLAVSDLLVLLGMPLEVEYEMRNYPFLFGVGCYFKT 120
 OY 124 ALFETVCFASILSTTVSVERYVAILHPFRAKLQSTRRRRLRLGLGWGFSVLSPLPTS 183
 DB 121 ALFETVCFASILSTTVSVERYVAILHPFRAKLQSTRRRRLRLGLGWGFSVLSPLPTS 180
 OY 184 IHGKFPNGSLVPGSATCTVIKPMWYINFTIOVTSFLFYLLPMTVISLVLYLMA 243
 DB 181 IHGKFPNGSLVPGSATCTVIKPMWYINFTIOVTSFLFYLLPMTVISLVLYLMA 240
 OY 244 KDKSLAEDEGNANIORPCRSYKMKLFVLVLFALICWAPFHIDRLFFSFVEEMSESLAA 303
 DB 241 KDKSLAEDEGNANIORPCRSYKMKLFVLVLFALICWAPFHIDRLFFSFVEEMSESLAA 300
 OY 304 VFNLVHVSGVFYLSAVNPITYNLLSRFOAFQNVISSEFKQWHSQDLPOLPPAQRN 363
 DB 301 VFNLVHVSGVFYLSAVNPITYNLLSRFOAFQNVISSEFKQWHSQDLPOLPPAQRN 360
 OY 364 IFTECHFEVLETDIGPOFPCOSSMHNHSLPTALSSQMSRTYQSFHFNKT 415
 DB 361 IFTECHFEVLETDIGPOFPCOSSMHNHSLPTALSSQMSRTYQSFHFNKT 412

RESULT 3

O9JIB1 PRELIMINARY: PRT: 395 AA.

AC 09JIB1
 ID 09JIB1
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE NEUROMEDIN U RECEPTOR 2.
 GN NMUR2R.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE=20351041; PubMed=10894543;
 RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
 RA Zeng Z., Williams D.L., Feigheimer S.D., Nunes C.N., Murphy B.,
 RA Stair J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McKee K.K.,
 RA Hrenluk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austin C.P.,
 RA Caskey T., van der Ploeg L.H.T., Liu Q.;
 RT "Identification of receptors for neuromedin U and its role in
 feeding.";
 RL Nature 406:70-74(2000).
 DR EMBL: AF242875; AAF82756.1; -;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCR_RHODPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECCEP_FL_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECCEP_FL_2; 1.

KW Receptor. 395 AA; 44722 MW; 01D3765B5D5355C0 CRC64;
SQ SEQUENCE

Query Match 75.5%; Score 1650.5; DB 11; Length 395;
Best Local Similarity 79.4%; Pred. No. 9.4e-115;
Matches 312; Conservative 36; Mismatches 40; Indels 5; Gaps 1;

```

OY 4 MEKLNASWIIYQOKLEDDPQKHLSNTEEYLAFICGPRSHFFLPVSVVYPIFVGVIGN 63
DB 1 MGLKENASWIIH-----DPLMKYLNSTEEYLAHLGCPKRSDSLPSVAVALITLVGMGN 55
OY 64 VLVCVLILQHOAKMKTPTNNYLFSLAVSDLLVLLGMPLEVEYEMRNYPFLFGVGCYFKT 123
DB 56 LVCVYIVRHQTKPTNNYLFSLAVSDLLVLLGMPLEIYEMNHNYPFLFGVGCYFKT 115
OY 124 ALPEYVCFASILSTTVSVRYVAIIHPRAKQSTRRALRLILGIWGSVLSPLPNTS 183
DB 116 ALPEYVCFASILSTTVSVRYVAIIHPRAKLESTRRALRLILSLWSESVFSLPNTS 175
OY 184 IHGIFKHPNGSLVPGSACTVYKPMIYNFIQVTSFLEYLIPMTVISLVLYMALRL 243
DB 176 IHGIFKHPNGSSVPGSACTVYKPMIYNFIQVTSFLEYLIPMTVISLVLYMALRL 235
OY 244 KKDSLEADGEGNANIORPCRSYVNMKLFVLVFAICMAFPHIDRLFFSFVEESESIAA 303
DB 236 KRDESLKANKVAVNIHPRSKSVTKMLFVLVFAICMFPFHIDRLFFSFVEESESIAA 295
OY 304 VENVLVHVSQVFYFYLSSAVNPITYNLSRFOAFQNVISFHKQHSQHDQDLPAPQNR 363
DB 296 VENVLVHVSQVFYFYLSSAVNPITYNLSRFRRAFRNVSPCKWCHPRHQPGPAPQKI 355
OY 364 IFLTECHFEVLTEDDIGPOPCOSSMHNHSLPTA 396
DB 356 IFLTECHFEVLTEDDIGPOPCOSSMHNHSLPTA 388

RESULT 4
OQESQ4 PRELIMINARY: PRT: 395 AA.
AC OQESQ4:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE G PROTEIN-COUPLED RECEPTOR TGR-1.
GN TGR-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=10887190;
RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
RA Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
RA Fujino M.;
RT "Identification and Functional Characterization of a Novel Subtype of
RT Neurexin Receptor."
RL J. Biol. Chem. 275:29528-29532(2000).
DR EMBL; AB041239; BAB13722.1;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1.1;
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G_PROTEIN_RECPEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECPEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 395 AA; 44756 MW; EABD6D36E9C355DA CRC64;
```

Query Match 75.2%; Score 1643.5; DB 11; Length 395;
Best Local Similarity 78.9%; Pred. No. 3.1e-114;
Matches 310; Conservative 37; Mismatches 41; Indels 5; Gaps 1;

```

OY 4 MEKLNASWIIYQOKLEDDPQKHLSNTEEYLAFICGPRSHFFLPVSVVYPIFVGVIGN 63
DB 1 MGLKENASWIIH-----DPLMKYLNSTEEYLAHLGCPKRSDSLPSVAVALITLVGMGN 55
OY 64 VLVCVLILQHOAKMKTPTNNYLFSLAVSDLLVLLGMPLEVEYEMRNYPFLFGVGCYFKT 123
DB 56 LVCVYIVRHQTKPTNNYLFSLAVSDLLVLLGMPLEIYEMNHNYPFLFGVGCYFKT 115
OY 124 ALPEYVCFASILSTTVSVRYVAIIHPRAKQSTRRALRLILGIWGSVLSPLPNTS 183
DB 116 ALPEYVCFASILSTTVSVRYVAIIHPRAKLESTRRALRLILSLWSESVFSLPNTS 175
OY 184 IHGIFKHPNGSLVPGSACTVYKPMIYNFIQVTSFLEYLIPMTVISLVLYMALRL 243
DB 176 IHGIFKHPNGSSVPGSACTVYKPMIYNFIQVTSFLEYLIPMTVISLVLYMALRL 235
OY 244 KKDSLEADGEGNANIORPCRSYVNMKLFVLVFAICMAFPHIDRLFFSFVEESESIAA 303
DB 236 KRDESLKANKVAVNIHPRSKSVTKMLFVLVFAICMFPFHIDRLFFSFVEESESIAA 295
OY 304 VENVLVHVSQVFYFYLSSAVNPITYNLSRFOAFQNVISFHKQHSQHDQDLPAPQNR 363
DB 296 VENVLVHVSQVFYFYLSSAVNPITYNLSRFRRAFRNVSPCKWCHPRHQPGPAPQKI 355
OY 364 IFLTECHFEVLTEDDIGPOPCOSSMHNHSLPTA 396
DB 356 IFLTECHFEVLTEDDIGPOPCOSSMHNHSLPTA 388

RESULT 5
OQ3664 PRELIMINARY: PRT: 403 AA.
AC OQ3664:
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ORPHAN G PROTEIN-COUPLED RECEPTOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tan C.T., McKee K.K., Palyha O.C., Feighner S.D., Hreniuk D.L.,
RA Smith R.G., Howard A.D.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF044601; AAC02680.1;
DR EMBL; AF044600; AAC02680.1; JOINED.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PR00001; 7tm.1.1;
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G_PROTEIN_RECPEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECPEP_F1_2; 1.
SQ SEQUENCE 403 AA; 44980 MW; CF8D5F1A961CCBD9 CRC64;
```

Query Match 43.1%; Score 942; DB 4; Length 403;
Best Local Similarity 50.8%; Pred. No. 2.9e-62;
Matches 180; Conservative 63; Mismatches 89; Indels 22; Gaps 6;

```

OY 20 DPOKHLSNTEEYLAFICGPRSHFFLPVSVVYPIFVGVIGNVLCVILQHOAKMKT 78
DB 13 DP--EDNLNLTDEALRLKYLGPQOTELPMPICATYLLFVVGAVNGTCLVILRHKMKRT 70
OY 79 PTNYLFLSLAVSDLLVLLGMPLEVEYEMRNYPFLFGVGCYFKTALFETVCFASILST 138
DB 71 PTNYLFLSLAVSDLLVLLGMPLEIYEMNHNYPFLFGVGCYFKTLLFEVYCLASVLTNT 130
OY 139 TVSVRYVAIIHPRAKQSTRRALRLILGIWGSVLSPLPNTSHGIFKHPNGSLV 198
DB 131 ALSVERVAVNHPLQARSMVTRAHVRYLGAVGLMCLSPNTSLHGIFQDLHVPGRGPV 190
OY 199 PGSATCTVYKPMIYNFIQVTSFLEYLIPMTVISLVLYMALRLKKDSLEADGEGN--- 255
```

```

Db 191 PSASVCMVLRPRALYNNVQTALLFFCLPMALMSVLYLLIGLRRLRRRLLLMQEAKGRG 250
OY 256 ----ANTQRC-----RKSVMKMLFVLVYVAICMAPFHIDRLFFSFYEESSESLAAV 304
Db 251 SAARSRYTCRLQOHDGRGRQVTKMLFVLVVGICMAPFHADRVMSVVSQMTDGLHLA 310
OY 305 FNLVHVSGVFYLSAVNPDIYNILSRFOAFQNVY---SSFH--KQWHSQH 353
Db 311 FOHVHVISGIFFLGSAANPVLVLSMSRFRFQALCLGACCHLRPRHSH 364

RESULT 6
O9HB89 PRELIMINARY: PRT: 426 AA.
ID O9HB89:
AC O9HB89:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE NEUROMEDIN U RECEPTOR 1.
GN NMUR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=20490668; PubMed=10899166;
RA Raddatz R., Wilson A.E., Artyushyn R., Bonini J.A., Borowsky B.,
RA Boteju L.W., Zhou S., Kouranova E.V., Nagorny R., Guevarra M.S.,
RA Dal M., Lerman G.S., Vayase P.J., Branchek T.A., Gerald C., Fortay C.,
RA Adham N.;
RT "Identification and Characterization of Two Neuremodin U Receptors
RT Differentially Expressed in Peripheral Tissues and the Central Nervous
RT System.";
RL J. Biol. Chem. 275:32452-32459(2000).
DR EMBL: AF272362; AAG24793.1; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1.1.
DR PRINTS: PR00237; GPCR_RHODOPSN.
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR_FL1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECEPTOR_FL2; 1.
KW Receptor.
SQ SEQUENCE 426 AA: 47350 MW: F8DD781C020F04AA CRC64;

```

```

Query Match 43.1%; Score 942; DB 4; Length 426;
Best Local Similarity 50.8%; Pred. No. 3.1e-62;
Matches 180; Conservative 63; Mismatches 89; Indels 22; Gaps 6;

```

```

OY 20 DFOKHLNSTEYLAF-LGPRSRHFPLPSVYVYPIFVGVGNVLCVLIHQHAKMT 78
Db 36 DF--EDLNTLDEALRLKYLGGPOTELFMPICATYLLIFVGVANGNLCTCVILRHKAMRT 93
OY 79 PNNYLFSLASDGLVLLGLMPLVEYEMKRNPFLEPGVGCYKTLAFTVFCASLISIT 138
Db 94 PNNYLFSLASDGLVLLGLMPLVEYEMKRNPFLEPGVGCYKTLAFTVFCASLISIT 153
OY 139 TVSEYEVVAIILPFRKLOSTRRALRIIGIVGFSVLSPTSIHGKIFENYPSGSLV 198
Db 154 ALSVEYEVVAIILPFRKLOSTRRALRIIGIVGFSVLSPTSIHGKIFENYPSGSLV 213
OY 199 PGSACTVTKPMWYVNIIVTSFLLYLPMTVISVLYLMLALRLKDKSLDEAGN--- 255
Db 214 PSASVCMVLRPRALYNNVQTALLFFCLPMALMSVLYLLIGLRRLRRRLLLMQEAKGRG 273
OY 256 ----ANTQRC-----RKSVMKMLFVLVYVAICMAPFHIDRLFFSFYEESSESLAAV 304
Db 274 SAARSRYTCRLQOHDGRGRQVTKMLFVLVVGICMAPFHADRVMSVVSQMTDGLHLA 333
OY 305 FNLVHVSGVFYLSAVNPDIYNILSRFOAFQNVY---SSFH--KQWHSQH 353
Db 334 FOHVHVISGIFFLGSAANPVLVLSMSRFRFQALCLGACCHLRPRHSH 387

```

```

RESULT 7
ID O9JIB2 PRELIMINARY: PRT: 402 AA.
AC O9JIB2:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE NEUROMEDIN U RECEPTOR 1.
GN NMUR1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=20351041; PubMed=10894543;
RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
RA Zeng Z., Williams D.L., Feighner S.D., Nunes C.N., Murphy B.,
RA Stair J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McKee K.K.,
RA Hreniuk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austin C.P.,
RA Caskey T., van der Ploeg L.H.T., Liu Q.;
RT "Identification of receptors for neuremodin U and its role in
RT feeding.";
RL Nature 406:70-74(2000).
DR EMBL: AF242873; AAF82754.1; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1.1.
DR PRINTS: PR00237; GPCR_RHODOPSN.
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR_FL1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECEPTOR_FL2; 1.
KW Receptor.
SQ SEQUENCE 402 AA: 45782 MW: CB576083D9F77A095 CRC64;

```

```

Query Match 40.5%; Score 885.5; DB 11; Length 402;
Best Local Similarity 48.3%; Pred. No. 4.5e-58;
Matches 173; Conservative 56; Mismatches 108; Indels 21; Gaps 4;

```

```

OY 19 EDPQKNI-----LNSTEEYLAFLGCPRRSHFPLPSVYVYPIFVGVGNVLCVLIHQ 72
Db 5 DSEKHEFDLEDNLTHEDRLKYLGGPQVQKFLCVTLLIFVGVANGNLCTCVILR 64
OY 73 HOAKMPTNYLFSLASDGLVLLGLMPLVEYEMKRNPFLEPGVGCYKTLAFTVFCASL 132
Db 65 QKAMHPTNYLFSLASDGLVLLGLMPLVEYEMKRNPFLEPGVGCYKTLAFTVFCASL 124
OY 133 SILSTTVSEYEVVAIILPFRKLOSTRRALRIIGIVGFSVLSPTSIHGKIFENY 192
Db 125 SVLNVTALSVERYVAIILPFRKLOSTRRALRIIGIVGFSVLSPTSIHGKIFENY 184
OY 193 PMSGLVPGSATCTVYKRMWYVNIIVTSFLLYLPMTVISVLYLMLALRLKDKSL 252
Db 185 PCGRPVADSVCTVYKRMWYVNIIVTSFLLYLPMTVISVLYLMLALRLKDKSL 244
OY 253 EGNANIQRPC-----RKSVMKMLFVLVYVAICMAPFHIDRLFFSFYEES 298
Db 245 EVKGRISAARQASHRSIQLRERROYTKMLALYVFGTCWVPFHADRLMSVMSHWT 304
OY 299 ESLAAVFNLVHVSGVFYLSAVNPDIYNILSRFOAFQNVY---SSFH--KQWHSQH 356
Db 305 DGLRLAFQSVHLASGVFLYLSAANPVLVLSMSRFRFQALCLGACCHLRPRHSH 361

RESULT 8
ID O9JUI5 PRELIMINARY: PRT: 412 AA.
AC O9JUI5:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

```


RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mlshina N.V., Mobery C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL: AEO03703; AAF55016.1; -;
 DR FlyBase: FBgn0038201; CG9918.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCRHHODPSN.
 DR PROSITE: PS00237; G-PROTEIN_RECPT_FL1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECPT_FL2; 1.
 KW G-protein coupled receptor; Glycoprotein; Transmembrane.
 SQ SEQUENCE 428 AA: 47198 MW: 9DDA41E930468BD7 CRC64;

Query Match 25.3%; Score 552.5; DB 5; Length 428;
 Best local similarity 33.2%; Pred. No. 2, 2e-33;
 Matches 121; Conservative 73; Mismatches 98; Indels 73; Gaps 7;
 38 GPRRS--HFLPVSVYVPIFVGVGNVLCVLIHQAMKPTNYLFLSLAVSDLLV 95
 11 GPRDPLAIVIPYVYVSLFIFGVGNISTCIVIKRSMHATNYLFLSLAISDFLL 70
 96 LILCMPLLEVEMNRYPRFLPGVCYFETALFETPCFASISITTVSERYVAILHPRAK 155
 71 LSGVPOEVSTIWSKYPRVEFEYICIGGLAETSANNTVLTITAFYERITAIHPPLGQ 130
 156 LQSTRRLRLIIGVGSVLSLPNTSINGIFHFYFPNGSLVPGSATCTVIRPMIYNF 215
 131 AMSKLSRAIRIIVLWIMAVTAIPQMAQIGIE-HY-----SGVQGCIVIRYVHSHF 182
 216 IIVTSTSLFLPLPVYISVLYIYIMALKDKSLDEADGANIQRPCR----- 263
 183 -QLSTFIIFELAPMSITLVLLIGVHLRSTLV---EGPASYARRQOLKSVSDTLVLR 237
 264 -----KSVNKLMLVLVLVAICW 281
 238 YGSGSTAMNSNGGSGAGTGLMGSGAQLSVRGRLNHYGTRVLMLAAVAVVCEFLCW 297
 282 APFHIDRLFPSEVSESESL-----AAVFNLVHVSGVFYLSAVNPFIYNLSSRFQAA 337
 298 APFHAQRLIAIVAPARGAKRQDHEFYVTVYVSGVLYLSTCINPLNLNINSHKREAA 357
 338 FQVNT 342
 358 FKAVL 362
 RESULT 11
 Q9VFW6 PRELIMINARY; PRT; 595 AA.

AC Q9VFW6; 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CG8795 PROTEIN.
 GN CG8795.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Gale R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blaziel R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Frankkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mlshina N.V., Mobery C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL: AEO03699; AAF54929.1; -;
 DR FlyBase: FBgn0038139; CG8795.
 DR InterPro: IPR000923; GPCR_blue1.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCRHHODPSN.
 DR PROSITE: PS00196; COPPER_BLUE; UNKNOWN_1.
 DR PROSITE: PS00237; G-PROTEIN_RECPT_FL1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECPT_FL2; 1.
 KW G-protein coupled receptor; Glycoprotein; Transmembrane.
 SQ SEQUENCE 595 AA: 65091 MW: 3EBE183BADP5471E CRC64;

Query Match 23.6%; Score 515; DB 5; Length 595;
 Best local similarity 35.6%; Pred. No. 1, 8e-30;
 Matches 110; Conservative 66; Mismatches 95; Indels 38; Gaps 9;
 QY 48 VSVYVYVPIFVGVGNVLCVLIHQAMKPTNYLFLSLAVSDLLVLLGLMPLLEVEM 107


```

Db 65 LSVGALFIFAGVGLNLTICIVISRNNEHMTATNFYLFALISDMILLCGMPQDLYNLM 124
Oy 108 R--NRPFLGPGCYCFKRALFETVCFASLTSTTVSVERVAILHPFRAKLOSTRRAIR 165
Db 125 HPDNP--FSDSICILESVLSTAAANAVLTITTAFTVYIAICHFROHTSKLSRAVK 182
Oy 166 ILGIWGFVSFLPSLNTSIGHKIFHPNGSLVPSATCTVAKPMIYFIQVTSFLPY 225
Db 183 FFAIWMALALLAP-----QAIOFSVMQGM-----CTSTGMNDFFAH--VFAVSGLFF 232
Oy 226 LIPMTVISLVYLMALRLKDKSLADEGNANIQRPC-----RKSVMKLFVLYLV 276
Db 233 GGPMTAICVLYLVIGVKIKRSRLQA-----LPRRCYDVNRGISAQTRIVIMLVAVAVA 286
Oy 277 FAICAPAFIDRLFFSF-----VEEMSESLAAVFNLVHVSQVFYLSAANPIIYNL 330
Db 287 FETICAPAFIAQRLAVYSGTSGIESQWEND--VSILDYTGVLVFLSTCINPLLYNIM 343
Oy 331 SRRFOAFAQ 339
Db 344 SHKFEAFK 352

RESULT 12
OyVFW5 PRELIMINARY: PRT: 660 AA.
AC 09VFW5:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CG8784 PROTEIN.
GN CG8784.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.C., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Botchva D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrelia S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris T.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,
RA Jajelli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Maltel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klimos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spredling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

```

```

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: AE003699; AAF54930.1; -.
DR FlyBase: FBgn0038140; CG8784.
DR InterPro: IPR000923; Copper_blue1.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PROSITE: PS00196; COPPER_BLUE; UNKNOWN_1.
DR PROSITE: PS00237; G_PROTEIN_RECPT_FL_1; 1.
DR PROSITE: PS50262; G_PROTEIN_RECPT_FL_2; 1.
DR G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ SEQUENCE 660 AA; 72277 MW; C52DIEAT6ACD439 CRC64;

Query Match 22.8%; Score 497.5; DB 5; Length 660.
Best Local Similarity 29.6%; Pred. No. 4e-29;
Matches 133; Conservative 77; Mismatches 159; Indels 81; Gaps 14;

Oy 20 DPEOKH----LNSTBEYLAFLCG-----PARRHFL--PVSYVYVIFVWG 59
Db 62 DKFLTHVAHLNITTEENSLNLTSGNGTNASTMAADSPVDESILTRTALTVCAVALIFVAG 121
Oy 60 VIGNVLCVLIIOHOAKPTNYLYFLSLAVSDLVLLGLMPLEVEEMHNP--FLFGPV 117
Db 122 VLGNLITCIVISRNNEHMTATNFYLFALISDMILLCGMPQDLYNLM--YDWMPEFDA 179
Oy 118 GGYFTALFETVCFASLTSTTVSVERVAILHPFRAKLOSTRRAIRLIGVWGSFYLE 177
Db 180 MCIMSGVLESEMANATVLTITTAFTVYIAICHFROHTSKLSRAIKRIFAIMLAFL 239
Oy 178 SLPNSTHIGIKHPNGSLVPSATCTVAKPMIYFIQVTSFLPYLIPMTVISLVLY 237
Db 240 ALPQAMQFSVVYQ-----NEGYSCTMEND--FYAHFAVSGFLFEGPMTAICVLYV 289
Oy 238 LMAFLRKDK--SLEADGNANIQRPCRSYKMKLFVLYLVFAICWAPFHIDRLFFSF- 293
Db 290 LIGVKIKRSRLQSLPRFIDANRGLMAGRIYIMLVAVAVAFELCMAPFHQRLMAVYG 349
Oy 294 ----VEEMSESLAAVFNLVHVSQVFYLSAANPIIYNLSSRRFOAFAQNT----- 342
Db 350 LNLINIGISRDAFNDFRILDYTGVLVFLSTCINPLLYNIMSHKFEAFKITLRQFGL 409
Oy 343 --SFFHKQ-WHGOHD-----POLPPA-----QRNIFLTCHEVELTE 376
Db 410 ANNHQSQOHQHNHNSALLRONGSMRLDPASCVSNNNALEPYGYRVQFRC-----R 463
Oy 377 DIGQPFCCSMHNSHLPALSSQMSRTN 406
Db 464 DANHQLSDSDIRTTTTTTTTTINSNSMAAGN 493

RESULT 13
OyVFW5 PRELIMINARY: PRT: 378 AA.
AC 018701:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE SIMILAR TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. NCBI GI: 1055105.
GN C48C5.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabdilitoidea;
OC Rhabdilitidae; Pelodertinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]

```

RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE-94150718; PubMed-7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstein L., Wilkinson-Sproat J., Wohlman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Favello A.;
 RL Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.
 DR EMBL: U39994; AAB37017.1; -;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm.1. 1
 DR PRINTS: PR00237; GPCRHHODPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
 DR SEQUENCE 378 AA; 43631 MW; 458DC60CBF8F86 CRC64;

Query Match 22.0%; Score 480; DB 5; Length 378;
 Best Local Similarity 31.8%; Pred. No. 4,6e-28;
 Matches 118; Conservative 67; Mismatches 146; Indels 40; Gaps 8;
 Oy 23 QKHLNTEEEY---LAFLCGRSH-----FLLPVSVVVP 56
 Db 3 QACLNTEEOCCCLAFNCPVYSHSEKACMEHCPSKRALDDVTLKATYALIFIF 62
 Oy 57 VVGIVGNVLVCIYLIOHQAANKPTNYLFLSLAVSDLLVLLGMPLEVEKMRNYPFLFGP 116
 Db 63 LVGVIGNTTCTCLVMKKHPMKTHASMYLMNLAVSDLVLCVGLPEVMMNMQYPPFPD 122
 Oy 117 VGCYFTALFEYVCFASITTSITVSERYVALIHP--FRAKLOSTRRALRILGIVGFSV 175
 Db 123 YICNLALALAEETSSVSLITLILFALERYVAVCHPLEKVKOPFKRNIGITIGFTWFSI 182
 Oy 176 LFSLPMTSHIGIKF--HYFP---NGSLVPGSATCTVI-----KPMIYNPIIOVTSFLFY 225
 Db 183 LCMAPRAIHHRADYIMKSNPGTDNRIPVSKSKMCAVMEPEFLATFKILHFSALAF 242
 Oy 226 LRPMTIVSVLYLMLRLKDKSLEADEGNANIOPRCRSVNMKLEFVLVLAFCAPRH 285
 Db 243 ALPLFTIVILARIACKVSSNRTIQ--PGELDITEELOHNRINAILCAIVSAFICVLPQ 300
 Oy 286 IDRLFFSPVEENSELAAPNLVHVYVGVFFYSSAVNPITVLLSRROAQAQNVISSE 345
 Db 301 LQRLFFYFDN--EVLITWVNOQYWFYISGLEFLATIIINPLAVNLASRRERAKDILIDY 359
 Oy 346 HKOMHSDHPQ 356
 Db 360 CMRGSGSERIPR 370

RESULT 14
 ID 017239 PRELIMINARY; PRT: 418 AA.
 AC 017239;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE K1084.4 PROTEIN.
 GN K1084.4
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita; Rhabditoidea;
 OC Rhabdilitidae; Peloderinae; Caenorhabditis.
 OC NCBI_Taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE-94150718; PubMed-7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstein L., Wilkinson-Sproat J., Wohlman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Sammons L., Wohlmann P., Antoniou B.;
 RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF025463; AAB71009.1; -;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm.1. 2.
 DR PRINTS: PR00237; GPCRHHODPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
 DR SEQUENCE 418 AA; 47735 MW; CE0416539CA3B27 CRC64;

Query Match 21.5%; Score 470.5; DB 5; Length 418;
 Best Local Similarity 31.8%; Pred. No. 2,6e-27;
 Matches 115; Conservative 76; Mismatches 130; Indels 41; Gaps 9;
 Oy 25 HLNSTEEYLAFLCGR--RSHEFLPVSVVYPIEVVGVGNVLVCIYLIOHQAANKPTNY 82
 Db 9 NVSEITEYVLTSGERCQAGIIVPIIYIGTFLILGICNTCIVIAANKSMHNPVY 68
 Oy 83 YLFSLAVSDLLVLLGMPLEVE--MMRNYPFLGPGVCFKRALPFTVCFASITTSVTS 141
 Db 69 YLFSLAVSDILILIGLPMFEYOSLDYSPYRSEGCARAFLEFTSYASIMITCPS 128
 Oy 142 VERRYVALHPFRAKLOSTRRALRILGIVGFSVLFSLPNTSHIGI----- 187
 Db 129 FERWLAICHPKSKIFSTLMKRVNLIILMTISFVCAIPIAFIVQINKLPDEDAKQYQW 188
 Oy 188 --KHFYF---NGSLVPGSAT-----CTVIRPMIYNEIIQVTSF--LEYLLPMTV 231
 Db 189 TNKVSFEAVGVLLNRIPEVTDGIFVLHTEFCAMNQSRDQOKMIIIFAFVFEVIPA 248
 Oy 232 ISVLYLMLALRLKDK--SLEADEGNANIOPRCRSVNMKLEFVLVLAFCAPFHIDLF 290
 Db 249 IYIMYAHIAVQLESSEIDKGDGMVKRRKRNKSNRTVLKMLLSVITTFICWLPFHQRL- 307

OY 291 FSEVENESES-----LAAVENLVHVGVEFFYSSAVNPITYNLSRRFOAQNVISS 344
 Db 308 LSVYTTMSETTTISPPVOFLSMIVFYISGFCYSSNSAANPILYNLSQKYSACFCRTILG 367
 OY 345 FH 346
 Db 368 DH 369

RESULT 15

O93413 PRELIMINARY; PRT; 363 AA.
 AC O93413;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ORPHAN G PROTEIN-COUPLED RECEPTOR.
 OS Spheroides nepheius.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes;
 CC Tetraodontidae; Spheroides.
 CX NCBI_TaxID=39110;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Palyha O.C., Tan C., McKee K.K., Feighner S.D., Hreniuk D.L.,
 RA Smith R.G., Howard A.D.;
 RT "Orphan G protein-coupled receptor from teleost fish Spheroides
 nepheius related to growth hormone secretagogue receptor."
 RL Submitted (AUG-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF082210; AAC33473.1;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS: PR00237; GPCR_RHODPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECPEP_F1_1; UNKNOWN_1.
 DR PROSITE: PS0262; G_PROTEIN_RECPEP_F1_2; 1.
 DR SEQUENCE 363 AA; 41324 MM; 74518BD9CD310991 CRC64;
 SQ

Query Match

Best Local Similarity 21.5%; Score 469; DB 13; Length 363;
 Matches 111; Conservative 64; Mismatches 102; Indels 26; Gaps 8;

OY 46 LPVSYYVPIFVGVYIGNVLVLCVLIQHOAMKPTNYLFSIAVSDLVLLGMPLEVEY 105
 Db 39 IPVYICILIFVGVGTGNTMILIIQYFKDKMTTNLVLSMAVSD-LVIFLCLEPDLXR 97
 OY 106 MKRNPPLFEGPYCYFKTALFETVCFASILSTTVSVERYVAIHPRAKLOSTRRALR 165
 Db 98 LMKYVPMLEGEAVCRLYHYIFEGCTSATILHTALSIERYLAISPELRKVMVTRRRVQY 157
 OY 166 ILGIWGSVLESLPNTSIHGKIFHFPGSLVPSATCT--VIRPMIYNFIIOVTSFL 223
 Db 158 ILLALMCFALVSAAPTLPLVGVENETHPDYNTGCKHGTGAISSGOLHIMVSTYF 217
 OY 224 FYLLPMTVIVLYLMLALRLKKKSLDEAGNANIQRPC-----RKSVMKLFVLVL 275
 Db 218 F--CPMLCLLFYGSIGCKLMSKN-----DLQPCALARESHRQTV-KILVVVL 266
 OY 276 VFAICWAPHIDRLFSEVENESES-LAAVENLVHVGVEFFYSSAVNPITYNLSRRF 334
 Db 267 AFIIICMLPYHIGRNLFADVDYDTAMLSQNFNM--ASMYLCYLSASINPVYVYINLSRKY 323
 OY 335 OAA 337
 Db 324 RAA 326

Search completed: April 22, 2002, 14:19:18
 Job time: 336 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:12:47 ; Search time 21.56 Seconds
(without alignments)
433.157 Million cell updates/sec

Title: US-09-609-146-4
Perfect score: 2185
Sequence: 1 MSGMEKLNASMTYQKLED.....ALSSSEOMRTMYGSHFNKT 415

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/pdata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/pdata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/pdata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/pdata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/pdata/2/1aa/6C.COMB.pep:*
6: /cgn2_6/pdata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	448	20.5	353	4	US-09-077-675A-3 Sequence 3, Appl1
2	444	20.3	361	4	US-09-077-675A-8 Sequence 8, Appl1
3	444	20.3	366	4	US-09-077-675A-13 Sequence 13, Appl1
4	436.5	20.0	364	4	US-09-077-675A-16 Sequence 16, Appl1
5	417.5	19.1	353	1	US-08-118-270-45 Sequence 45, Appl1
6	417.5	19.1	353	5	PCT-US93-08528-45 Sequence 45, Appl1
7	410	18.8	416	3	US-08-858-876A-4 Sequence 4, Appl1
8	410	18.8	416	4	US-09-472-880-4 Sequence 4, Appl1
9	405.5	18.6	398	2	US-08-288-663A-1 Sequence 2, Appl1
10	402	18.4	302	4	US-09-077-675A-2 Sequence 2, Appl1
11	398	18.2	302	4	US-09-077-675A-7 Sequence 7, Appl1
12	398	18.2	410	3	US-08-858-876A-2 Sequence 2, Appl1
13	398	18.2	410	4	US-09-472-880-2 Sequence 2, Appl1
14	397	18.2	393	1	US-07-629-1041-3 Sequence 3, Appl1
15	377.5	17.3	398	2	US-08-288-663A-15 Sequence 15, Appl1
16	365	16.7	369	1	US-07-816-283-6 Sequence 6, Appl1
17	365	16.7	369	1	US-08-417-103-6 Sequence 6, Appl1
18	365	16.7	369	1	US-08-417-103-16 Sequence 16, Appl1
19	358	16.4	369	2	US-08-411-859-3 Sequence 9, Appl1
20	358	16.4	369	4	US-08-387-707-9 Sequence 9, Appl1
21	357	16.3	400	4	US-08-188-275A-2 Sequence 2, Appl1
22	356.5	16.3	319	4	US-08-832-399-2 Sequence 2, Appl1
23	356.5	16.3	319	4	US-09-372-498-2 Sequence 2, Appl1
24	356	16.3	369	4	US-08-120-601B-9 Sequence 9, Appl1
25	355	16.2	400	3	US-08-889-108-8 Sequence 8, Appl1
26	355	16.2	400	5	PCT-US94-10358-8 Sequence 8, Appl1
27	352.5	16.1	380	3	US-08-676-351-5 Sequence 5, Appl1

28	352.5	16.1	380	4	US-08-188-275A-5 Sequence 5, Appl1
29	351	16.1	384	3	US-09-071-434-3 Sequence 3, Appl1
30	351	16.1	389	4	US-08-430-286A-7 Sequence 7, Appl1
31	350.5	16.0	376	4	US-08-387-707-17 Sequence 17, Appl1
32	350.5	16.0	380	4	US-08-765-743-2 Sequence 2, Appl1
33	349	16.0	369	1	US-07-816-283-8 Sequence 8, Appl1
34	349	16.0	369	1	US-08-417-103-8 Sequence 8, Appl1
35	348.5	15.9	380	1	US-08-149-093A-7 Sequence 7, Appl1
36	348.5	15.9	380	2	US-08-911-245-7 Sequence 2, Appl1
37	348.5	15.9	380	2	US-08-147-592A-2 Sequence 2, Appl1
38	348.5	15.9	380	4	US-09-510-473-7 Sequence 7, Appl1
39	347.5	15.9	391	1	US-07-816-283-2 Sequence 2, Appl1
40	347.5	15.9	391	1	US-07-816-283-4 Sequence 4, Appl1
41	347.5	15.9	391	1	US-08-417-103-2 Sequence 2, Appl1
42	347.5	15.9	391	1	US-08-417-103-4 Sequence 4, Appl1
43	347.5	15.9	391	1	US-08-417-103-14 Sequence 14, Appl1
44	345.5	15.8	367	3	US-08-676-351-2 Sequence 2, Appl1
45	345.5	15.8	391	4	US-08-120-601B-8 Sequence 8, Appl1

ALIGNMENTS

RESULT 1
US-09-077-675A-3
Sequence 3, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feigner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077, 675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 353 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-077-675A-3
Query Match 20.5%; Score 448; DB 4; Length 353;

RESULT 6
 PCT-US93-08528-45
 : Sequence 45, Application PC/TUS9308528
 : GENERAL INFORMATION:
 : APPLICANT: New York University
 : TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
 : TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
 : NUMBER OF SEQUENCES: 348
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: BROWDY AND NEIMARK
 : STREET: 419 Seventh Street, N.W., Suite 300
 : City: Washington
 : STATE: D.C.
 : COUNTRY: USA
 : Zip: 20004
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: PCT/US93/08528
 : FILING DATE: 09-SEP-1993
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/943,236
 : FILING DATE: 10-SEP-1992
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Townsend, Kevin G.

RESULT
US-08-858-876A-4
Sequence 4, Application US/08859876A
Patent No. 6022856
GENERAL INFORMATION:
APPLICANT: Daniel CAPUT
APPLICANT: Pascale CHALON
APPLICANT: Pascual FERRARA
APPLICANT: Vita NATALIO
TITLE OF INVENTION: Type 2 Neurotensin Receptor
TITLE OF INVENTION: (hNT-R2)
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
STREET: 400 Seventh Street
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/858,876A
FILING DATE: 19-SEP-1997
CLASSIFICATION: 536

RESULT 8
 US-09-472-880-4
 Sequence 4, Application US/09472880
 Patent No. 6274333
 GENERAL INFORMATION:
 APPLICANT: Daniel CAPUT
 Pascale CHALON
 Pascual FERRARA
 Vito NATALIO
 TITLE OF INVENTION: Type 2 Neurotensin Receptor
 (hnt-R2)
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
 STREET: 400 Seventh Street
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/472,880
 FILING DATE: 28-Dec-1999
 CLASSIFICATION: <unknown>
 PRIOR APPLICATION DATA:

RESULT 9
 US-08-288-663A-1
 : Sequence 1, Application US/08288663A
 : Patent No. 5879896
 : GENERAL INFORMATION:
 : APPLICANT: HINDMA, Shuji
 : APPLICANT: HOSODA, Masaki
 : APPLICANT: ONDA, Haruo
 : TITLE OF INVENTION: HUMAN TRH RECEPTOR, ITS PRODUCTION
 : TITLE OF INVENTION: AND USE
 : NUMBER OF SEQUENCES: 15
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
 : STREET: 130 Water Street
 : CITY: Boston
 : STATE: MA
 : COUNTRY: USA
 : ZIP: 02109
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FASTSEQ for Windows Version 2.0
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/288, 663A
 : FILING DATE: 09-AUG-1994
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 198309/1993

FILING DATE: 10-AUG-1993
APPLICATION NUMBER: 286986/1993
FILING DATE: 16-NOV-1993
APPLICATION NUMBER: 325215/1993
FILING DATE: 22-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 44612
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400.
TELEFAX: 617-523-6440
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TOPOLOGY: 1 linear
MOLECULE TYPE: protein
US-08-288-663A-1

Query Match 18.6%; Score 405.5; DB 2; Length 398;
Best Local Similarity 29.9%; Pred. No. 2.1e-30;
Matches 119; Conservative 72; Mismatches 130; Indels 77; Gaps 14;

QY 48 VSVVVPPIPV-VGVIGNVLVCLVLIHQAMKTPNTYLSLAVSDLLVL-LGMP-L--E 102
DB 27 VVLLVLLICGLGIVGINVIMVLMVMTKHMRTPTNCLVSLAVADLVLAAGLPNTIDS 86
QY 103 VEMAMNHYFELFCGVCYEFALFETVCFASILITTVSVERVAIIHPRAKLOSTRR 162
DB 87 IYGSW-----VGYVGCCLCTITYLOYIGINASSCSITAFIERIAICHPKAOFCTFSR 141
QY 163 ALRILGIWGSVLSFSLPNTSHGIKFHFPGSLVPGSATC-----TVIKPMWITNFI 216
DB 142 AKKILFVVAFTSLYCMLEFLLDLNISTYKDAIVI-----SCGYKISRNYSPYILMDFG 197
QY 217 IQVTSFLFLLPMTVTSVLYYLALRL-----KKKSLDADGNNINQ 260
DB 198 V-----FVYVPHILATVLLGFARILFLPISDPKRENSKTKNDSTHONTNLNVTNS 231
QY 261 PC-----RKSVMKMLFVLLVFAICWAFH---IDRLFFS---FVEEMSESLAAVFNLY 308
DB 252 RCNSTSVSSRKQYTKMLAVVILFALLMMPYRLVYVNSFLSPQENM-----F 301
QY 309 HVSQGFYFLSSAVNPITYNLSRRQAFAQNVISFHKOMHSQHPOLPRAORNIFLT- 367
DB 302 LFCRICIYLNSAIVNVIYNLSQKFRFAFRKLCNCOK-----PTKRPANYSVALNY 354
QY 368 -----ECHFVELTEDIGQPPCCSSMHNLSLPTALSSE 400
DB 355 SVIKESDHSTELDDITVTDYLSATKVSFDDTCLASE 392

RESULT 10

US-09-077-675A-2
Sequence 2, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pal, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA

ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 302 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: 1 linear
MOLECULE TYPE: protein
US-09-077-675A-2

Query Match 18.4%; Score 402; DB 4; Length 302;
Best Local Similarity 33.6%; Pred. No. 3.2e-30;
Matches 95; Conservative 64; Mismatches 96; Indels 28; Gaps 7;

QY 68 LVLIHQAMKTPNTYLSLAVSDLLVL-LGMPLEVEEMRNYPFLFGPGCYFKTALFE 127
DB 2 LVVSRREMTNTNLVLSMAFSDDLIFLC-MPLDLFRLMQYRPMNLGNLCKLQFVSE 60
QY 128 TVCFASLSTTVSVERVAIIHPRAKLOSTRRLRIIGIWSGVLSLPTNTSHGT 187
DB 61 SCTYATVLLITLALSVREYFAICPLRAKVVTGKVKLVLYLVAVAFCAGPIFVLGV 120
QY 188 KHFYFPGS-----LVPGSATCVIKPMWITNFIQVTSFLFLLPMTVJSYL 235
DB 121 E---HDNGIDPDTNCRATEFAVRSGLLTVM---VWV-----SVFFFLVFCITVL 167
QY 236 YVLMALRLKKDKSLDEGNANIQRCKRSVMKMLFVLLVFAICWAFHIDRLFFS-FV 294
DB 168 YSLIGRKLMMRRKGEAAVG-SLRDQNHKQTVKMLAVVFAFLCMLPFIHGRYLFSSKL 226
QY 295 EEMSESLAAVFNLYHVSQGFYFLSSAVNPITYNLSRRQA 337
DB 227 EPGSVEIAQISQYCNLSFVLYLSAIVNLSIMSKKYRA 269

RESULT 11

US-09-077-675A-7
Sequence 7, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pal, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ

```

STATE: D. C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/858, 876A
FILING DATE: 19-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 9723204
FILING DATE: 17-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Player, William E.
REGISTRATION NUMBER: 31,049
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 410 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-858-876A-2
Query Match 18.2%; Score 398; DB 3; Length 410;
Best Local Similarity 29.2%; Pred. NO. 1,1e-29;
Matches 106; Conservative 77; Mismatches 118; Indels 62; Gaps 13;

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

; ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
; STREET: 400 Seventh Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/472,880
; FILING DATE: 28-Dec-1999
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR 9723204
; FILING DATE: 17-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Player, William E.
; REGISTRATION NUMBER: 31,049
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-472-880-2

```

```

Query Match      18.2%; Score 398; DB 4; Length 410;
Best Local Similarity 29.2%; Pred. No. 1.1e-29;
Matches 106; Conservative 77; Mismatches 118; Indels 62; Gaps 13;

```

```

QY 49 SVYVPTGVGVGNVLVCLVLIQHOMAKT-PTNYVLFSLAVSDLVLLIGMLEVYE-M 106
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 36 TALYALTMALGAAGNMLSVHVLKARAGRLRHVLSLAGLLLLLVGVPELTSFV 95
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 107 WRVYVPLFGVPGCG---YFKALFETVCFASILSTTVSVRYVALIHPRAKIOSTPRRA 163
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 96 WFNYPWFGDLGRGYF---VHELCAIVATVLSVAGISAERCLAVCOPLRARSILPRRT 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 164 LRLIGVWGFVSLESLPNTSINGIKFHY-FPNGSLVPGSATCTVYIKPMYINFIQVTSF 222
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 153 RMVLASMAASGLALPMAYIMGQKHELETADEPERPASVGVIVLSRTALQVFIQVNL 212
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 223 LFTLLPRTVLSVL-----YLMAL-----RLKKDSLEADG----- 234
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 213 VSFVLLPLALTAFLNGVTVSHILALCSQVPSSTPGSSPSRLK-----LISEEGILSPITW 268
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 255 -NANIO-----RPCRKSVMKMLFVLVLFVFAICWAPFHIDLFPSFV--EE 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 269 KRFPIGGGOVSLVRHKDVRIRSLQNSV-OVLAIIVMYICWLPYHARLMALCYPPDA 327
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 297 WSESLAIVNLVHVVSQVFYLLSSAVNPITVNLSSRFFQAFQNVISSEFKHOMHQHPQ 356
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 328 WTPPLNFEYHYFYWMTLTFLYVSSAVTPLYLVNVSSEFRKLFLKAVSSSLGCHHPM--KR 385
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 357 LPP 359
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 386 LPP 388

```

RESULT 14
US-07-629-1041-3

```

; Sequence 3, Application US/076291041
; Patent No. 5288621
; GENERAL INFORMATION:
; APPLICANT: Gershengorn, Marvin C
; APPLICANT: Straub, Richard E
; TITLE OF INVENTION: PITUITARY TRH RECEPTOR
; NUMBER OF SEQUENCES: 3

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 3.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/629,1041
; FILING DATE: 19901214
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; REFERENCE/DOCKET NUMBER: CRF D - 995
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-629-1041-3

```

```

Query Match      18.2%; Score 397; DB 1; Length 393;
Best Local Similarity 28.6%; Pred. No. 1.3e-29;
Matches 110; Conservative 72; Mismatches 104; Indels 98; Gaps 15;

```

```

QY 18 LEDPFOKHLNSTE-----EY-----LAFLCGPRSRSHFELPYSVYVPIPVGVYI 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 MENDTVSENMQTELOPQAAVALEFQVVTLLVITIGC-----LGIV 41
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 GNVLVCLVLIQHOMAKTPTNYVLFSLAVSDLVLL-LGML---EYEMKRNTPFLGPIV 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 42 GNMVVLVWRTKMRPTNCYLVSLAVADLVLAAGLEPNIDTSYGSW-----YGYV 96
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 GCVTKALFETVCFASILSTTVSVRYVALIHPRAKIOSTRRLRLIGIVWGFVSIF 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 97 GCICITYLYQLGINASSCSITTAFTIERIYAIQPIKAQFLCTSRAKIITIFWAFSTIY 156
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 178 SLPNTSINGIKFHYFPNGSLVPGSATC-----TVIKPMYINFIQVTSFLLPMTV 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 157 CMLMFLLDINISYKNKAVV-----SCGYKISRNYVSPILMDGV-----FYVPMIL 206
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 232 ISVLYTYMA-----LRLKKDSLEADG--NANIORC-----RKSVMK 268
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 207 ATVLGFIARILFLNPISDPKNSKMKNDSIHONKLNLMNTNCFNSTVSSRKQVTK 266
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 269 MLFVLVLFVFAICWAPFH---IDRLPFS--FVEEMSESLAIVNLVHVVSQVFYLLSAVN 323
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 267 MLAVVILFALMLMPTLTVLVNVSFLSSPQENM-----FLFCICITLYLSAIN 316
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 324 PIYVNLSSRRFOAFQNVISSEFKH 347
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 317 PVYINLMSQKFAAFRLCKCKOK 340

```

RESULT 15
US-08-288-663A-15

```

; Sequence 15, Application US/08288663A
; Patent No. 5879896
; GENERAL INFORMATION:
; APPLICANT: HINUMA, Shuji
; APPLICANT: HOSoya, Masaki
; APPLICANT: ONDA, Haruo

```


...

FT	LIPID	388	388	PALMITATE (POTENTIAL).
SO	SEQUENCE	424 AA;	47054 MM;	AAC2ETFEAF8DBCD3 CAC6d:
	Query Match	22.4%;	Score 489.5;	DB 1; Length 424;
	Best Local Similarity	32.3%;	Pred. No. 1.6e-23;	
	Matches 107;	Conservative 75;	Mismatches 114;	Gaps 35; Indels 9;
OY	48	VSVVVPVFPEVVGCGIGNVLVCYLVILOH---	QAMKTPPNYYLSLAVSDLVLLGLMPLEYV	104
		: : : : : : : : : : : : : : :	: : : : : : : : : : : :	
Db	67	VTATLAEFVGVGTNSVTAFTLARKKSLOSIVNHLGSLADSLDILLAPRVELY		126
OY	105	E-MMRNYEPFLCEGPVC---YFKTALFEFYCFASISITTVSVERVAIILHPRAKLOSR		160
Db	127	NFIWNNHHMWAGODACCRCGYF---LRDACTVATALNVASLSVERKLALCHPRKAATLMNR		183
OY	161	RRALRLICIGVMGSVLESPLNPSTSHGIKFHYPFGNSLVPGSATCTVIKPMWIYNELIOYT		220
Db	184	SRTKKFISAIWIASLALLAIIPLMFTGLONR--SGDGTHRGGLVCPVIDTAVKKVIQVN		241
OY	221	SFLFVLPMPTVSVLYUUMALARLKDKDSLEADEG-----NANIQRPCKR		264
Db	242	TFMSPFLFPLMLYSILTNTVIANKLIVYMHOAEQGRCVGTGHNGLEHSTFNNTIE-PGRV		300
OY	265	SVNK----MLFVLYLVFAICMAFFHIDLRFESFV--EENSSESLAVFNLVHVVGCVFFL		318
Db	301	QALRHGVALVLRVVIAVFAVYCMLPVHYVRBLMFCYISDEQWTTELPFGYHNHFVLTNALFPV		360
OY	319	SSAVNPPIIYNLNSRRFOQAFOINVISSFRKOW		349
Db	361	SSAINPIDLYNLVSANFRQVFLSTLACICDPGW		391

RESULT	2			
NRRL_MOUSE		STANDARD:	PRT:	424 AA.
NR1_MOUSE				
OC	088319;			
DT	20-AUG-2001 (Rel. 40, Created)			
DT	20-AUG-2001 (Rel. 40, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	NEUROTENSIN RECEPTOR type 1 (NT-R-1).			
GN	NTSR1 OR NTSR.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RA	Spider J., Sano H., Ohba M.;			
RT	"Neurotensin receptor type 1.";			
RL	Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS			
CC	ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-			
CC	CALCIN SECOND MESSENGER SYSTEM.			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.			
CC	HIGHEST TO TACHIKININS RECEPTORS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/).			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AB017027; BAA33013.1; -			
DR	MCD; MGI:97386; Ntsr.			
DR	InterPro; IPR000276; GPCR_Rhodpsn.			
DR	Pfam; PF00001; 7tm_1.1.			
DR	PRINTS; PR00237; GPCR_RHODPSN.			
DR	PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; 1.			

DR	PROSITE, PSS0262; G-PROTEIN_RECIP_F1_2; 1.	KW	G-protein coupled receptor; Transmembrane; Glycoprotein;
RN	Phosphorylation; Lipoprotein; Palmitate.	RN	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN 1	FT	1 (POTENTIAL).
FT	TRANSMEM 65	FT	1 (POTENTIAL).
FT	DOMAIN 88	FT	2 (POTENTIAL).
FT	TRANSMEM 97	FT	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN 122	FT	3 (POTENTIAL).
FT	TRANSMEM 144	FT	3 (POTENTIAL).
FT	DOMAIN 166	FT	4 (POTENTIAL).
FT	TRANSMEM 189	FT	4 (POTENTIAL).
FT	DOMAIN 211	FT	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 235	FT	5 (POTENTIAL).
FT	TRANSMEM 260	FT	5 (POTENTIAL).
FT	DOMAIN 308	FT	6 (POTENTIAL).
FT	TRANSMEM 309	FT	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN 331	FT	7 (POTENTIAL).
FT	TRANSMEM 349	FT	7 (POTENTIAL).
FT	DOMAIN 373	FT	8 (POTENTIAL).
FT	CARBOHYD 4	FT	9 (POTENTIAL).
FT	CARBOHYD 38	FT	10 (POTENTIAL).
FT	CARBOHYD 42	FT	11 (POTENTIAL).
FT	CARBOHYD 211	FT	12 (POTENTIAL).
FT	DISULFD 141	FT	13 (POTENTIAL).
FT	LIPID 388	FT	14 (POTENTIAL).
SO	SEQUENCE 424 AA; 47216 MW; 8E9A723171A48711 CRC64;	SO	15 (POTENTIAL).

Db	36	LSNCGSNSESLPEPNSLNDVNTDITYSKVL-----VTAVYLLAFVYGVIGNSVT	84
Qy	67	CLVLIQH---QAMKTPTYVYLSLAVSDLVLLCMPELVEY-MWRNPYLEGPGVC---	119
Db	85	AFLLARKKSLQSLQSTVYHYHGLSLASDLILLMLPVELYNFIVWHHPMAFGDAGCGY	144
Qy	120	YFKTALFETVCFASLTSTTVSVERYVAIILHPEKAKLOSTRRLRLILGIVGFSVLSL	179
Db	145	YF---LRDACTAYATALNVAASLSEVERYLAIChPEKAKTLMSSRSRTKKFTSLVIALSALLAV	201
Qy	180	PRTSHHGKGFHFHPGSL---VPGSATCTVYIKPMIYVFIQVTSFELYLLPMYISLV	236
Db	202	PMLEFTMGLO---NRSADGQHGGSLVCTPTVDATATVYVQVNTFMSEFLPMLIISILN	256
Qy	237	YLMALRK-----KDRSLSEADGEGNANIORPCRSYVK---MLFVYL	275
Db	257	TVIAKKLVYVYVHQAAEOGRGVCVTGTHNSLEHSTFNMSE-PCRQALRHGVLYLRAVAY	315
Qy	276	VFAICMAPEFHIDRLFEEFV---EEMSESLAAVFNLYVVSQGVFFYLSAVNPITYMLLSRR	333
Db	316	AFVYCMLEHYHVRRLAFGVCISDQWTFLEDFHYHYMTLNLFLYVSSAINPILNYLSAN	375
Qy	334	FOAFQNVVISSEFHQW 349	
Db	376	FROVFLSTLACLPCGW 391	
RESULT 3			
GP38_HUMAN	GP38_HUMAN	STANDARD;	PRT; 412 AA.
AC	043193:		
DT	15-DEC-1998 (Rel. 37, Created)		
DT	15-DEC-1998 (Rel. 37, Last sequence update)		
DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	PUTATIVE G PROTEIN-COUPLED RECEPTOR GP38.		
GN	GP38.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
NCBI_TaxId	9606;		


```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98110578; PubMed=9441746;
RA McKee K.K., Tan C.P., Palyna O.C., Liu J., Feighner S.D.,
RA Hreniuk D.L., Smith R.G., Howard A.D., van der Ploeg L.H.T.;
RT "Cloning and characterization of two human G protein-coupled receptor
RT genes (GPR38 and GPR39) related to the growth hormone secretagogue
RT and neurensin receptors.";
RL Genomics 46:426-434(1997).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN THYROID, STOMACH, AND BONE
CC MARROW.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF034632; AAC26081.1; -.
DR GCRDB; GCR_2494; -.
DR MIM; 602885; -.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 2.
DR PRINTS: PR00237; GPCR_RHODPSN.
DR PROSITE: PS00237; G_PROTEIN_REC_F1_1; 1.
DR PROSITE: PS00262; G_PROTEIN_REC_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 35 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 36 56 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 57 74 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 75 94 2 (POTENTIAL).
FT DOMAIN 95 112 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 113 134 3 (POTENTIAL).
FT DOMAIN 135 157 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 158 178 4 (POTENTIAL).
FT DOMAIN 179 246 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 247 270 5 (POTENTIAL).
FT DOMAIN 271 298 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 299 320 6 (POTENTIAL).
FT DOMAIN 321 334 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 335 358 7 (POTENTIAL).
FT DOMAIN 359 412 CYTOPLASMIC (POTENTIAL).
FT DISULFD 111 235 BY SIMILARITY.
FT CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 412 AA; 45344 MW; C13FF6165012DEF3 CRC64;

```

```

Query Match 21.8%; Score 476; DB 1; Length 412;
Best Local Similarity 31.8%; Pred. No. 1e-22;
Matches 115; Conservative 81; Mismatches 106; Indels 60; Gaps 11;

DB 266 LVGLIGREILMSRRPLRGPAASGRERGHQTV-RVLLVVLAFIICMLPFHGRIIYNT 324
OY 295 EEMSESL--AAFNVLHVVSQVFYLLSSAVNPITVLLSRRQ-AAFOVVIS-----SF 345
DB 325 ED-SRMVFSQYENIVAL--QLFYLSASINPLVNLISKYRAAFKLLARKSRPGE 380
OY 346 HK 347
DB 381 HR 382

RESULT 4
ID NTIR1_HUMAN STANDARD; PRT; 418 AA.
AC P30989;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NEURENIN RECEPTOR TYPE 1 (NT-R-1) (HIGH-AFFINITY LEVOCABASTINE-
DE INSNSITIVE NEURENIN RECEPTOR) (NTRH).
OS NTIR1 OR NTRR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93154505; PubMed=8381365;
RA Vita N., Laurent P., Lefort S., Chalon P., Dumont X., Kaghad M.,
RA Gully D., Le Fur G., Ferrara P., Caput D.;
RT "Cloning and expression of a complementary DNA encoding a high
RT affinity human neurensin receptor.";
RL FEBS Lett. 317:139-142(1993).
CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEURENIN. IT IS
CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIN SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHIKININ RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X70070; CAA49675.1; -.
DR PIR: S29506; S29506.
DR GCRDB; GCR_0577; -.
DR MIM; 162651; -.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PROSITE: PS00237; G_PROTEIN_REC_F1_1; 1.
DR PROSITE: PS00262; G_PROTEIN_REC_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 63 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 64 86 1 (POTENTIAL).
FT DOMAIN 87 95 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 96 120 2 (POTENTIAL).
FT DOMAIN 121 142 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 143 164 3 (POTENTIAL).
FT DOMAIN 165 187 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 188 209 4 (POTENTIAL).
FT DOMAIN 210 234 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 235 259 5 (POTENTIAL).
FT DOMAIN 260 303 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 304 325 6 (POTENTIAL).
FT DOMAIN 326 343 EXTRACELLULAR (POTENTIAL).

```

FT TRANSMEM 344 367 7 (POTENTIAL).
 FT DOMAIN 368 418 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 37 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 141 224 BY SIMILARITY.
 FT LIPID 383 383 PALMITATE (POTENTIAL).
 SQ SEQUENCE 418 AA; 46288 MW; BBB1EEC2BE6390 CRC64;

Query Match 21.6%; Score 471; DB 1; Length 418;
 Best Local Similarity 31.0%; Pred. No. 2.1e-22;
 Matches 120; Conservative 72; Mismatches 121; Indels 74; Gaps 14;

20 DPQKHLNSTEEL-----AFLCGRRS-----HFLPVSVVYVPEV 58
 17 DPFQROAGLEALLPAGFNGASGSEKYLAPSSSLDVNTDYSKVLVTANYALFV 76
 59 GVIGNVLVCLVILQH---QAMKTPYNYLFSLAVSDLLVLLGMPLEVE-MNRNPELF 114
 77 GTVGNVTATFTLARKKSLQSLQSVHYHGLSLASDLTLTLLAMPVELYVFWHHWAF 136
 115 GPVGC---YKTLAFETVCASLSTTVERVVALHPFRAKLOSTRRLRLIGIV 171
 137 GDAGCGRQYF---LRDAGTATALNVAASLSVERYLAIHPFKAKTMSRSRRTKFTSAIW 193
 172 GFGLVSLP-----NTSIGHIKFHPNGSLV--PGSATCTYIKPMIYNFIQVTS 221
 194 LASALLTVPMLFTWGEONKRSADQH---AGGLVCTPTIHTIV-----KVIOVNT 241
 222 FLEYLEPMTVISVLYLMLRLKDKSLDEGN-----ANIQPRCK 264
 242 FMSFIPPMVIVISVLTNTIAKLTVMVQAEGQVCGEHSFESALIPRGVQ-ALRH 300
 265 SVKMLFVLVLAICAPRHIDLPFSEV--EEMSESLAVNLVHVSQVFFYSSAV 322
 301 G--RVLRVAVIAVVCMLPRVHRLMFCYISDQWTPFLDYHFYMTNALFVYSTI 359
 323 NPITYNLSSRFOAFONVISFHKW 349
 360 NPILYNVSNFRHIFLATIACLPW 386

RESULT 5
 GHSR_PIG STANDARD: PRT: 366 AA.

AC 095254: 095255: PRT: 366 AA.
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GROWTH HORMONE SECRETAGOGUE RECEPTOR TYPE 1 (GHS-R) (GH-RELEASING PEPTIDE RECEPTOR) (GHRP).
 GN GHSR.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_Taxid=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-YORKSHIRE; TISSUE-Pituitary;
 RX MEDLINE=96337998; PubMed=8688086;
 RA Howard A.D., Feighner S.D., Cully D.F., Arena J.P., Liberator P.A., Rosenblum C.I., Hamelin M., Hreniuk D.L., Palyha O.C., Anderson J., Parass P.S., Diaz C., Chou M., Liu K.K., McKee K.K., Pong S.-S., Chaung L.-Y., Elbrecht A., Daskevitz M., Heavens R., Rigby M., Strinchingji D.J.S., Dean D.C., Mellilo D.G., Patcheit J.M., Nargund R., Griffin P.R., Demartino J.A., Gupta S.K., Scheffert J.A., Smith R.G., van der Ploeg L.H.T.;
 RA A receptor in pituitary and hypothalamus that functions in growth hormone release.";
 RT Science 273:974-977(1996).
 RL -I- FUNCTION: RECEPTOR FOR GROWTH HORMONE RELEASING PEPTIDES (GHRP) AS WELL AS NON-PEPTIDE, LOW MOLECULAR WEIGHT SECRETAGOGUES (E.G. L-

CC 692,429, MK-0677). THIS RECEPTOR IS COUPLED TO G-ALPHA-11
 CC PROTEIN. BINDING POTENCY FOR THE SECRETAGOGUES IS IN THE ORDER:
 CC MK-0677 > GHRP-2 > GHRP-6.
 CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN
 CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1A (SHOWN HERE) AND 1B; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING. ISOFORM 1B APPEARS NOT TO BIND
 CC SECRETAGOGUES.
 CC -I- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS.
 CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U60178; AAC48630.1; -
 DR EMBL: U60180; AAC48631.1; -
 DR GCRDB: GCR_1532; -
 DR GCRDB: GCR_1538; -
 DR InterPro: IPR003905; GHS1_receptor.
 DR InterPro: IPR002276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCRHHODOPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECP_FL_1; 1.
 DR PROSITE: PS0262; G_PROTEIN_RECP_FL_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Alternative splicing.

FT DOMAIN 1 40
 FT TRANSMEM 41 66
 FT DOMAIN 67 72
 FT TRANSMEM 73 96
 FT DOMAIN 97 117
 FT TRANSMEM 118 139
 FT DOMAIN 140 162
 FT TRANSMEM 163 183
 FT DOMAIN 184 211
 FT TRANSMEM 212 235
 FT DOMAIN 236 263
 FT TRANSMEM 264 285
 FT DOMAIN 286 302
 FT TRANSMEM 303 326
 FT DOMAIN 327 366
 FT DISULFID 116 198
 FT CARBOHYD 13 13
 FT CARBOHYD 27 27
 FT VARSPLIC 266 289
 FT VARSPLIC 290 366
 SQ SEQUENCE 366 AA; 41194 MW; 2C850B3BEF61B701C CRC64;

Query Match 20.5%; Score 448; DB 1; Length 366;
 Best Local Similarity 34.3%; Pred. No. 4.8e-21;
 Matches 104; Conservative 68; Mismatches 103; Indels 28; Gaps 7;

48 VSVYVPIEVGYIGNVCLVILQHQAMKTPYNYLFSLAVSDLLVLLGMPLEVEYEW 107
 46 VTCVALLFVVGAGNGLNLTLLVVSRRERMTNTLSSMAFSLDILFLC-MPLDLEFLM 104
 108 RNPPLFGPGVGCYKFNALFETVCASLSTTVERVVALHPFRAKLOSTRRLRLITL 167
 105 QYRPMNLGNLCLKFQVSSCTYATVLTALSVETFAICPLKAKVYVTKGRKLVY 164
 168 GYVGCVSFLSLENTSIGHIKFHPNGS-----LVGSATCTYIKPMIYNF 215
 165 LVIMVAFCAGPIFLVAGE---HNGTDPRTDTCNEGRATEFAVRGGLTYM--VWV--- 216
 216 IIOVTSFLEFLVLPMTVISVLYLMLRLKDKSLDEGNANIQRPCRSVKNMFLVYL 275
 217 -----SSVEFLFVFLVLYSLIGRKLWRKRGFAVNG--SSLRDQNHKQTVKMLAVVF 270

QY 276 VEAICWAFHIDRLFFS-FVEEMSESLAANLVHVSQVFYLSAANPIIYLNLSRRF 334
 DB 271 AFTLCMLPHVGVRLFSKSPGSEVLEIAISOVCNLYSVFLYLSAANPIIYLNLSRRF 330
 QY 335 QAA 337
 DB 331 RVA 333
 RESULT 6
 GHSR_HUMAN STANDARD: PRT: 366 AA.
 AC 092847: 092848: (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 40, Last annotation update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE GROWTH HORMONE SECRETAGOGUE RECEPTOR TYPE 1 (GHS-R) (GH-RELEASING
 DE PEPTIDE RECEPTOR) (GHRP).
 GN GHSR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=pituitary;
 RX MEDLINE=96337998; PubMed=8688086;
 RA Howard A.D., Feigheimer S.D., Cully D.F., Arena J.P.,
 RA Lillabator P.A., Rosenblum C.I., Hamelin M., Hreniuk D.L.,
 RA Palyha O.C., Anderson J., Pares P.S., Diaz C., Chou M., Liu K.K.,
 RA McKee K.K., Pong S.-S., Chong L.-Y., Elbrecht A., Daskevich M.,
 RA Heavens R., Rigby M., Strimlingh D.J.S., Dean D.C., Mellillo D.C.,
 RA Patchett A.A., Nargund R., Griffin P.R., Demartino J.A., Gupta S.K.,
 RA Schaeffer J.M., Smith R.G., van der Ploeg L.H.T.;
 RT A receptor in pituitary and hypothalamus that functions in growth
 RT hormone release.*;
 RT Science 273:974-977(1996).
 RL -1- FUNCTION: RECEPTOR FOR GROWTH HORMONE RELEASING PEPTIDES (GHRP) AS
 CC WELL AS NON-PEPTIDE, LOW MOLECULAR WEIGHT SECRETAGOGUES (E.G. L-
 CC 692,429, MK-0677). THIS RECEPTOR IS COUPLED TO G-ALPHA-11
 CC PROTEINS. POTENCY RANKING FOR SECRETAGOGUE INHIBITION IS MK-0677
 CC > GHRP-2 > GHRP-6.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1A (SHOWN HERE) AND 1B: ARE
 CC PRODUCED BY ALTERNATIVE SPLICING. ISOFORM 1B APPEARS NOT TO BIND
 CC SECRETAGOGUES.
 CC -1- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U60179; AAC50653.1; -
 CC EMBL: U60181; AAC50654.1; -
 CC DR GCRDB: GCR_1917; -
 CC DR GCRDB: GCR_1918; -
 CC DR MIM: 601898; -
 CC DR InterPro: IPR003905; GHS1_rcptor.
 CC DR InterPro: IPR000276; GPCR_Rhodopsn.
 CC DR Pfam: PF00001; 7tm_1; 1.
 CC DR PRINTS: PR00237; GPCR_Rhodopsn.
 CC DR PROSITE: PS00237; G-PROTEIN_RECEPTOR_FL_1; 1.
 CC DR PROSITE: PS50262; G-PROTEIN_RECEPTOR_FL_2; 1.
 CC KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 CC Alternative splicing.
 FT DOMAIN 1 40
 FT TRANSMEM 41 66 1 (POTENTIAL).
 FT RT

FT DOMAIN 67 72
 FT TRANSMEM 73 96
 FT DOMAIN 97 117
 FT TRANSMEM 118 139
 FT TRANSMEM 140 162
 FT TRANSMEM 163 183
 FT TRANSMEM 184 211
 FT TRANSMEM 212 235
 FT TRANSMEM 236 263
 FT TRANSMEM 264 285
 FT TRANSMEM 286 302
 FT TRANSMEM 303 326
 FT TRANSMEM 327 366
 FT DISULFID 116 198
 FT CARBOHYD 13 13
 FT CARBOHYD 27 27
 FT VARSPLIC 266 289
 FT VARSPLIC 290 366
 FT SEQUENCE 366 AA; 41328 MW; D1B62710DA9DC0C6 CRC64;
 SO
 Query Match 20.3%; Score 444; DB 1; Length 366;
 Best Local Similarity 34.0%; Pred. No. 8, 3e-21;
 Matches 103; Conservative 68; Mismatches 104; Indels 28; Gaps 7;
 QY 48 VSVVYVPIPVGVYIGVNLVCLVILQHOAKTPNNYVLFSLAVSDLVLLGMPLEYEMW 107
 DB 46 VYATVCALEFVAGIAGNLTLMLVSRRELRTTNLVLSSAFSDLLIFLC-MPLDLVRLM 104
 QY 108 RNPFLFPGVGCYFKALFETVCFASILSTTVSEYVYAILHPFAKLOSTRRLRLI 167
 DB 105 QYRPNMFGDLCKLFPVFSCTYATVLTITLALSVRYAICPLPAKVYTKGRKVI 164
 QY 168 GIWVGFSLSPNTSIHGKIFHYFPNGS-----LVGSATCTVIKPMIYNE 215
 DB 165 FVIMAVAFCSAGPIFVLGVGEH-----NGTDPMDTNECRPEFAVRSGLLTVM---VWV--- 216
 QY 216 IIOVTFPLFLPMYIVISVYIYIMALKKLEDEGANIQRCKRSYVMFLVYL 275
 DB 217 -----SIFFLVFLVCLVLYLSLIGKRLRRRGDAVVG-ASLRDNNHOTVYMLAVVF 270
 QY 276 VEAICWAFHIDRLFFS-FVEEMSESLAANLVHVSQVFYLSAANPIIYLNLSRRF 334
 DB 271 AFTLCMLPHVGVRLFSKSPGSEVLEIAISOVCNLYSVFLYLSAANPIIYLNLSRRF 330
 QY 335 QAA 337
 DB 331 RVA 333
 RESULT 7
 GHSR_RAT STANDARD: PRT: 364 AA.
 AC 008725:
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GROWTH HORMONE SECRETAGOGUE RECEPTOR TYPE 1 (GHS-R) (GH-RELEASING
 DE PEPTIDE RECEPTOR) (GHRP).
 GN GHSR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=pituitary;
 RX MEDLINE=97246555; PubMed=9092793;
 RA McKee K.K., Palyha O.C., Feigheimer S.D., Hreniuk D.L., Tan C.,
 RA Phillips M.S., Smith R.G., der Ploeg L.H.T., Howard A.D.;
 RT Molecular analysis of rat pituitary and hypothalamic growth hormone
 RT secretagogue receptors.*;

```

RL Mol. Endocrinol. 11:415-423(1997).
RN [2]
RP SEQUENCE OF 1-240 FROM N.A.
RC STRAIN-WISTAR; TISSUE=Plutitary;
RX MEDLINE=98100386; PubMed=9437732;
RA Yokote R., Sato M., Matsubara S., Onye H., Ntlimi M., Murto K.,
RT Takehana J.,
RT "Molecular cloning and gene expression of growth hormone-releasing
RT peptide receptor in rat tissues.";
RL Peptides 19:15-20(1998).
CC -1- FUNCTION: RECEPTOR FOR GROWTH HORMONE RELEASING PEPTIDES (GHRP) AS
CC WELL AS NON-PEPTIDE, LOW MOLECULAR WEIGHT SECRETAGOGUES (E.G. L-
CC 692,429, MK-0677). THIS RECEPTOR IS COUPLED TO G-ALPHA-11
CC PROTEINS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL: U94321; A0C53156.1; .
DR EMBL; AB001982; BAA21777.1; ALT_INIT.
DR GCRDP; GCR_1383; .
DR InterPro: IPR0003905; GHS1_rcptor.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm1.1;
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PROSITE: PS00237; G-PROTEIN_RECPT_F1.1; 1.
DR PROSITE: PS02627; G-PROTEIN_RECPT_F1.2; 1.
KW G-protein coupled receptor; transmembrane; glycoprotein.
FT DOMAIN 1 40
FT TRANSNM 41 66
FT DOMAIN 67 72
FT TRANSNM 73 96
FT DOMAIN 97 117
FT TRANSNM 118 139
FT DOMAIN 140 162
FT TRANSNM 163 183
FT DOMAIN 184 211
FT TRANSNM 212 235
FT DOMAIN 236 263
FT TRANSNM 264 285
FT DOMAIN 286 302
FT TRANSNM 303 326
FT DOMAIN 327 364
FT DISULFD 115 197
FT CARBOHYD 13 13
FT CARBOHYD 26 26
FT SEQUENCE 364 AA; 40963 MW; DCBF559BE061EE9B CAC64;

```

[illegible][illegible]

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN RP SEQUENCE FROM N.A.
 RX MEDLINE=93371401; PubMed=8395824;
 RA Mære V., Karlsten H.E., Wright M.S., Lundell I., Fjeldheim K.,
 RA Gabrielsen O.S., Lærdammar D., Gautvik K.M.;
 RT "Molecular cloning of a functional human thyrotropin-releasing
 RT hormone receptor.";
 RL Biochem. Biophys. Res. Commun. 195:179-185(1993).
 RN RP SEQUENCE FROM N.A.
 RC TISSUE=Plutary;
 RX MEDLINE=93384596; PubMed=8396925;
 RA Yamada M., Monden T., Satoh T., Satoh N., Murekami M., Iriuchijima T.,
 RA Kakegawa T., Mori M.;
 RT "Plutary adenomas of patients with acromegaly express thyrotropin-
 RT releasing hormone receptor messenger RNA: cloning and functional
 RT expression of the human thyrotropin-releasing hormone receptor
 RT gene.";
 RL Biochem. Biophys. Res. Commun. 195:737-745(1993).
 RN RP SEQUENCE FROM N.A.
 RC TISSUE=Plutary;
 RX MEDLINE=94063224; PubMed=8243797;
 RA Duthie S.M., Taylor P.L., Anderson L., Cook J., Eldne K.A.;
 RT "Cloning and functional characterisation of the human TRH receptor.";
 RL Mol. Cell. Endocrinol. 95:R11-R15(1993).
 RN RP SEQUENCE FROM N.A.
 RC TISSUE=Plutary;
 RX MEDLINE=95002135; PubMed=7918619;
 RA Hinuma S., Hosoya M., Ogi K., Tanaka H., Nagai Y., Onda H.;
 RT "Molecular cloning and functional expression of a human thyrotropin-
 RT releasing hormone (TRH) receptor gene.";
 RL Biochim. Biophys. Acta 1219:251-259(1994).
 RN RP SEQUENCE FROM N.A.
 RC TISSUE=Plutary;
 RX MEDLINE=9635621; PubMed=8703031;
 RA Iwasaki T., Yamada M., Satoh T., Konaka S., Ren Y., Hashimoto K.,
 RA Konga H., Kato Y., Mori M.;
 RT "Genomic organization and promoter function of the human thyrotropin-
 RT releasing hormone receptor gene.";
 RL J. Biol. Chem. 271:22183-22188(1996).
 RN RP SEQUENCE OF 1-263 FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=99101143; PubMed=9886052;
 RA Mære V., Hovring P.I., Orstevik S., Frengen E., Rian E.,
 RA Velickovic Z., Murray-McInosh R.P., Gautvik K.M.;
 RT "Structural and functional organization of the gene encoding the
 RT human thyrotropin-releasing hormone receptor.";
 RL J. Neurochem. 72:40-50(1999).
 CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
 CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
 CC PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC -----
 DR EMBL: D16845: BA004120.1; -;
 DR EMBL: X75071: CA52965.1; -;
 DR EMBL: X72089: CA50979.1; -;
 DR EMBL: S75283: AAB32322.1; -;
 DR EMBL: S75281: AAB32322.1; JOINED.
 DR EMBL: D85376: BAA12796.1; -;
 DR EMBL: D85375: BAA12796.1; JOINED.

DR	EMBL: AJ011701; CA009746.1; -	
DR	PIR: JN0759; JN0759.	
DR	PIR: SA0682; SA0682.	
DR	PIR: JN0708; JN0708.	
DR	GCRRdb: GCR_0763; -	
DR	GCRRdb: GCR_0799; -	
DR	GCRRdb: GCR_0880; -	
DR	GCRRdb: GCR_1111; -	
DR	GCRRdb: GCR_1148; -	
DR	MIM: 188545; -	
DR	InterPro: IPR000276; GPCR_Rhodopsn.	
DR	Pfam: PF00001; 7tm_1; 1.	
DR	PRINTS: PR00237; GPCRHHOOPS.	
DR	PRINTS: PR00751; THYROLIBRN.	
DR	PROSITE: PS00237; G_PROTEIN_RECIP.F1.1; 1.	
DR	PROSITE: PS50262; G_PROTEIN_RECIP.F1.2; 1.	
KW	G-protein coupled receptor; Transmembrane; Glycoprotein.	
FT	DOMAIN	1 28
FT	TRANSMEM	29 51
FT	DOMAIN	52 61
FT	TRANSMEM	62 83
FT	DOMAIN	84 99
FT	TRANSMEM	100 121
FT	DOMAIN	122 144
FT	TRANSMEM	145 168
FT	DOMAIN	169 193
FT	TRANSMEM	194 215
FT	DOMAIN	216 266
FT	TRANSMEM	267 288
FT	DOMAIN	289 296
FT	TRANSMEM	297 319
FT	DOMAIN	320 398
FT	CARBOHYD	3 3
FT	CARBOHYD	10 10
EQ	SEQUENCE	398 AA: 45084 MW: 929035FE33D3E CRC64: N-LINKED (GLCNAC. . .) (POTENTIAL) N-LINKED (GLCNAC. . .) (POTENTIAL) FE920B5FE33D3E CRC64: (POTENTIAL)

Query Match	18.68;	Score 405.5;	DB 1;	Length 398;
Best Local Similarity	29.98;	Pred. No. 2e-18;		
Matches 119;	Conservative 72;	Mismatches 130;	Indels 77;	Gaps 14

ID	RESULT	NTFR2_HUMAN	STANDARD:	PR7:	410 AA.
QY	48	VSYYVPIFV-VGVIQNVLCVLTILHQAAMKPTNNYLFSLASDLYL-LGMDL---	E 102		
Db	27	VTILLVLLICGLGIVGNINVLVYVARKHMTPTNCIYLSLAVADMLVYLAAGLPNTDS	86		
QY	103	YEMKRNYPLEPGVGYEKTALFEYGCASILSITTVSVERVYALLHPRAKLQSTRRR	167		
Db	87	IYGSW----YGVGVGCITCYLYLGILGNASSCSITTAETIRYIALICHPIAQGLCFEFSR	144		
QY	163	ALRILGIWGFVSYLESPDTPNTSIHGKIFHPFNGSLVPSASCC-----TVYKKPMWYNFI	211		
Db	142	AKRIILFIYMATLSLXCMHFFLLDLNISTYKDALYI-----SCGKIKISNYSPIYLDYFG	197		
QY	217	IQVTSLEFLLPMTYIVISLVYLYMALRL-----KKDSLEADDEGNANIQR	267		
Db	198	V-----FYVYPMILATLVGFIRILFLNDIPSDPKENSKYMKNDSTHQTNTLVNTSN	251		
QY	261	PC-----RKSVNKKLEVLVLVPAICAPHR---IDRLFS--FVEEMSESLAAVENTLV	308		
Db	252	RCFNSTYSRRQVYKMLAAVYILFLALLMMPRLTVLVNVSFLSPFGQEW-----F	304		
QY	309	HVSGCVFFYLSASAVNPITYNLISLRFOAFQNVYSSFFHKOMHSDHQLDLPRAQRIEFLT-	367		
Db	302	LLFCRICIYLLISAINPVIYVNLMSQKFRFAAFKRLCNCRK-----PTEKPNYSVALNY	354		
QY	368	-----ECHFEVLTFEDIGQFPQCSGMHNSHLPTALSS-400			
Db	355	SVIKESDHFSTFELDDITVDTVYLSATKVSFDDTCLASE 392			

AC 095665;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE NEUROKININ RECEPTOR TYPE 2 (NT-R-2) (LEVOCABASTINE-SENSITIVE
 DE NEUROKININ RECEPTOR) (NTR2 RECEPTOR).
 GN NTR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99066919; PubMed=9851594;
 RA Vicia N., Oury-Donat F., Chalon P., Guillemot M., Kaghad M., Bachy A.,
 RA Thureyssen S., Garcia S., Poinot-Chazel C., Casellas P., Keane P.,
 RA Le Fur G., Maffrand J.P., Shoubrite P., Caput D., Ferrara P.;
 RT "Neurotensin is an antagonist of the human neurotensin NT2 receptor
 RT expressed in Chinese hamster ovary cells.";
 RL Eur. J. Pharmacol. 360:265-272(1998).
 CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
 CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHIKININS RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y10148; CAA71233.1;
 DR MIM: 605538;
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm1.1;
 DR PRINTS: PR00237; GPCR_Rhodopsn.
 DR PROSITE: PS00237; G-PROTEIN_RECPEP_F1_1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECPEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.
 FT DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 33 55 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 56 64 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 65 87 2 (POTENTIAL).
 FT DOMAIN 88 109 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 110 131 3 (POTENTIAL).
 FT DOMAIN 132 154 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 155 176 4 (POTENTIAL).
 FT DOMAIN 177 217 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 218 237 5 (POTENTIAL).
 FT DOMAIN 238 297 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 298 318 6 (POTENTIAL).
 FT DOMAIN 319 337 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 338 358 7 (POTENTIAL).
 FT DOMAIN 359 410 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 108 194 BY SIMILARITY.
 FT LIPID 377 377 PALMITATE (POTENTIAL).
 SQ SEQUENCE 410 AA: 45413 MW: 8634222BEL15FD66 CRC64;

Query Match 18.2%; Score 398; DB 1; Length 410;
 Best Local Similarity 29.2%; Pred. No. 6e-18;
 Matches 106; Conservative 77; Mismatches 118; Indels 62; Gaps 13;

OY 49 SVVYPIPIVYIGNVLCVLIQHQAMKT-PTNYLFSLSVLDLVLLGKPLEYVE-M 106
 DB 36 TALLVLLIMVAGNALSIVVVKARAGRGRHNVLSLALGLLLLVGVEVLEISFV 95
 OY 107 MNVYPLFPGVGC---YFXTALFETVCFASISITTVSERYVALIHPRAKIQSTRRR 163

DB 96 WFIYPMVFGDLGGRGYF---VHELCAVATVLSVAGLSMRCLAVCOPLRARSLLPRRT 152
 OY 164 LRLIGIVMGESVLESLPNTSINGIKFHY-FPNCGLVPGSATCTVIRPMYNTIOTYTS 222
 DB 153 RMLVALSMASLGLALPMVIMGOKHELETADPEPASPACVTLVSRRALQVFIQVMVL 212
 OY 223 LPLFLPMVTSVY----YVLMAL-----RLKKDLSLEADEG----- 254
 DB 213 VSFVLPALAFNLAVNGVTVSHLLALCSQVPSTSPGSSTSRLE----LSEGLLSFIYW 268
 OY 255 -NANIQ-----RPCRKSVNKMFLVLYVAICAPRIDRIFSFY--EE 296
 DB 269 KTFEIGGGVSLVRHKDVRIRISLQRSV-QVLAIVMYVICLPRHARLMCIYPDDA 327
 OY 297 WSESLAIVNLVHVGSGVFYSSAVNPITLYNLISRRFOAFONVTSFRKQMSHQDPO 356
 DB 328 WPDPLNFHYFMYVNTLFYSSAVTPPLLYNAVSSSRFLFLEAVSSLCGEHHPM--KR 385
 OY 357 LPP 359
 DB 386 LPP 388
 RESULT 12
 TRFR_MOUSE STANDARD; PRT; 393 AA.
 AC P21761;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE THYROTROPIN-RELEASING HORMONE RECEPTOR (TRH-R) (THYROLIBERIN
 DE RECEPTOR).
 GN TRHR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=pituitary;
 RX MEDLINE=91088548; PubMed=2175902;
 RA Straub R.E., Frech G.C., Joho R.H., Gershengorn M.C.;
 RT "Expression cloning of a cDNA encoding the mouse pituitary
 RT thyrotropin-releasing hormone receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9514-9518(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92381047; PubMed=1324930;
 RA Narayanan C.S., Fujimoto J., Geras-Raaka E., Gershengorn M.C.;
 RT "Regulation by thyrotropin-releasing hormone (TRH) of TRH receptor
 RT mRNA degradation in rat pituitary GH3 cells.";
 RL J. Biol. Chem. 267:17296-17303(1992).
 RN [3]
 RP SEQUENCE OF 332-393 FROM N.A.
 RC TISSUE=pituitary;
 RX MEDLINE=97013702; PubMed=9156522;
 RA Jones K.E., Brubaker J.H., Chin W.W.;
 RT "An alternative splice variant of the mouse TRH receptor mRNA is the
 RT major form expressed in the mouse pituitary gland.";
 RL J. Mol. Endocrinol. 16:197-204(1996).
 CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
 CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
 CC PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@lsb.slb.ch.
CC EMBL: M59811; AAA0480.1; -
DR EMBL: M94384; AAA0437.1; -
DR EMBL: L48936; AAA81559.1; -
DR PIR: A39251; A39251.
DR GCRDB: GCR_0099; -
DR GCRDB: GCR_1613; -
DR MGD: MGI:98824; Trh.
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00751; THYROLIBRIN.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 28
FT DOMAIN 1 (POTENTIAL).
FT DOMAIN 2 51
FT DOMAIN 3 61
FT DOMAIN 4 83
FT DOMAIN 5 99
FT DOMAIN 6 122
FT DOMAIN 7 144
FT DOMAIN 8 168
FT DOMAIN 9 193
FT DOMAIN 10 215
FT DOMAIN 11 266
FT DOMAIN 12 288
FT DOMAIN 13 296
FT DOMAIN 14 319
FT DOMAIN 15 339
FT DOMAIN 16 393
FT CARBOHYD 3 10
FT CARBOHYD N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 393 AA; 44559 MW; 8739B75D1A0FCFC CRC64;

Query Match 18.2%; Score 397; DB 1; Length 393;
Best Local Similarity 28.6%; Pred. No. 6.7e-18;
Matches 110; Conservative 72; Mismatches 104; Indels 98; Gaps 15;

QY 18 LEPDFQKHNSTE-----EY-----LAFICGRSHFFLPVSVYVPIFVGVY 61
DB 1 MENDTVSEMNQTELOPOAAVALEYOVTVLLVITIC-----LGIV 41
QY 62 GNVLCVLYILOHQAKPTNYVLFSLAVSDLLVL--LGMPL--EYEMRNYPFLGPV 117
DB 42 GNMVLLVVMVRKHMRTPCNYLVSLAVADLVLAAGLNDISIGSM-----YGYV 96
QY 118 GCYFKALPETYVCFASLSTITVSVERYVAILHPRAKLOSTRRALRIIGIYWGSEVLF 177
DB 97 GCCTIYLYOYLGINASSCSITAFETIERYVIAICHPKQFLCTFSRAKKIIFWAFPSIY 156
QY 178 SLPTNSIHGKIFHPNGSLVPGSANC-----TVIKPMIYVFITIOVTSFLFYLPMTV 231
DB 157 CMLMFELDLNISTYKNAVYV-----SCGYKISKNYISPTILMDGV-----FYVPMIL 206
QY 232 ISVLYVYLA-----LRLKDKSLEADG--NANTORPC-----RKSYNK 268
DB 207 ATVLGYFIARILFLNPSPDKENSKMKMNDISQKNLNLNTNRCFNSTVSSRKQVTK 266
QY 269 MFLVLLVLAICAAPH---IDRLFFS--FVEKSESALAVALNLHVSVVFYLLSAVN 323
DB 267 MALVVVILFALLMMPYRTLVVNSFLSSPQENW-----FLLCRICIYLNLSAIN 316
QY 324 PIYVNLISRFQAFQVNSPFK 347
DB 317 PVIYNLMSQKFAFRKLCNCKOK 340

RESULT 13
TRFR_SHEEP STANDARD: PRT: 398 AA.
AC Q28396;
DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE THYROTROPIN-RELEASING HORMONE RECEPTOR (TRH-R) (THYROLIBERIN
DE RECEPTOR).
GN TRH.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eultheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Pituitary.
RX MEDLINE=97200775; PubMed=9048604;
RA Bockmann J., Boeckers T.M., Winter C., Wittkowski W., Winterhoff H.,
RA Deufel T., Kreutz M.R.;
RT "Thyrotropin expression in hypophyseal pars tuberalis-specific cells
is 3,5,3'-trideoxythionine, thyrotropin-releasing hormone, and pit-1
independent.";
RL Endocrinology 138:1019-1028(1997).
CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
CC PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.lsb.slb.ch/announce/>
CC or send an email to license@lsb.slb.ch).
CC
DR EMBL: X95285; CA64606.1; -
DR GCRDB: GCR_1294; -
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00751; THYROLIBRIN.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 28
FT DOMAIN 2 51
FT DOMAIN 3 61
FT DOMAIN 4 83
FT DOMAIN 5 99
FT DOMAIN 6 122
FT DOMAIN 7 144
FT DOMAIN 8 168
FT DOMAIN 9 193
FT DOMAIN 10 215
FT DOMAIN 11 266
FT DOMAIN 12 288
FT DOMAIN 13 296
FT DOMAIN 14 319
FT DOMAIN 15 339
FT CARBOHYD 3 10
FT CARBOHYD N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 398 AA; 45088 MW; 375A311D3D2A61A CRC64;

Query Match 18.1%; Score 394.5; DB 1; Length 398;
Best Local Similarity 29.6%; Pred. No. 9.6e-18;
Matches 118; Conservative 74; Mismatches 129; Indels 77; Gaps 14;

QY 48 VSVVYVPIFY-VGVIGNVLCVLYILOHQAKPTNYVLFSLAVSDLLVL--LGMPL--E 102
DB 27 VTLLVLLIGLGIVGINVLLVMVRKHMRTPCNYLVSLAVADLVLAAGLNDISIGSM 86
QY 103 VYEMRNYPFLFGVGYCYFTALFETVYCFASLSTITVSVERYVAILHPRAKLOSTRRR 162
DB 87 IYGSW-----VGYVGLCTIYLYOYLGINASSCSITAFETIERYVIAICHPKQFLCTFSR 141

OY 163 ALRLIGVWGSFSLPNTSIGHIKFHYFNGSLVPGSATC-----TVIKPMIYFI 216
 DB 142 AKKIIIFWAFISICMLMFLLDINISYKDAIV-----SCGYKSRMYSPYIAMDVG 197
 OY 217 IOVTFLEFLPMTVIVSYLYMALRL-----KKDSLEADEN 255
 DB 198 V-----FYVPMILATVLYGFIARILFLSPIPSPKRENSNTWKNDSTHQNKNLSKTSN 251
 OY 256 A--NIQRCRKSVMKMLFVLVFAICNAPRH---IDRLFFS--FVEESESIAAVFNLY 308
 DB 252 RYFNSTVSSRKQVMTMLAVVLEFLMLMPYFTLVVNSFSSPQENW-----F 301
 OY 309 HVAAGVFVFLSSAVNPITIIYNLRSRFOAFQNVISFHKQMHSDHPQOLPPAQRNIFLT- 367
 DB 302 LIFCICILYLSAINPVITNMLSKQFRAFRRLCNCOKR-----PVEKPNYSVALNY 354
 OY 368 -----ECHFVELTEDIGFQPCOSSMHSNHLPTALSS 400
 DB 355 SVIKESDHFSTELDDITVTDVLYSATKVSFDDTCLASE 392
 RESULT 14
 TRFR_RAT STANDARD: PRT: 412 AA.
 AC 001717; 063948; 063949;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE THYROTROPIN-RELEASING HORMONE RECEPTOR (TRH-R) (THYROLIBERIN RECEPTOR).
 GN TRHR.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;
 RX 11
 RC TISSUE: FROM N.A.
 RX MEDLINE=92283212; PubMed=1317787;
 RA Zhao D., Yang J., Jones K.E., Gerald C., Suzuki Y., Hogan P.G., Chin W.W., Tashjian A.H. Jr.;
 RT "Molecular cloning of a complementary deoxyribonucleic acid encoding the thyrotropin-releasing hormone receptor and regulation of its messenger ribonucleic acid in rat GH cells.";
 RL Endocrinology 130:3529-3536(1992).
 RN 12
 RC TISSUE: FROM N.A.
 RX MEDLINE=92322017; PubMed=1377915;
 RA de la Pena P., Delgado L.M., del Camino D., Barros F.;
 RT "Cloning and expression of the thyrotropin-releasing hormone receptor from GH3 rat anterior pituitary cells.";
 RL Biochem. J. 284:891-899(1992).
 RN 13
 RC TISSUE: FROM N.A.
 RX MEDLINE=9324585; PubMed=8387312;
 RA Sellar R.E., Taylor P.L., Lamb R.F., Zabavnik J., Anderson L., Eligne K.A.;
 RT "Functional expression and molecular characterization of the thyrotropin-releasing hormone receptor from the rat anterior pituitary gland.";
 RL J. Mol. Endocrinol. 10:199-206(1993).
 RN 14
 RC TISSUE: FROM N.A.
 RX STRAIN=SPRAGUE-DAWLEY; TISSUE=pituitary;
 RX MEDLINE=93246029; PubMed=1338727;
 RA Eligne K.A., Sellar R.E., Couper G., Anderson L., Taylor P.L.;
 RT "Molecular cloning and characterization of the rat pituitary gonadotropin-releasing hormone (GnRH) receptor.";
 RL Mol. Cell. Endocrinol. 90:R5-R9(1992).
 RN 15

RP SEQUENCE FROM N.A.
 RA MEDLINE=94102223; PubMed=8275956;
 RX Kimura N., Arai K., Sahara Y., Suzuki H., Kimura N.;
 RT "Estradiol transcriptionally and posttranscriptionally up-regulates thyrotropin-releasing hormone receptor messenger ribonucleic acid in rat pituitary cells.";
 RL Endocrinology 134:432-440(1994).
 RN 16
 RC TISSUE: FROM N.A. (ISOFORMS TRH-R(412) AND TRH-R(387)).
 RX MEDLINE=94158585; PubMed=7509436;
 RA Satoh T., Feng P., Wilber J.F.;
 RT "A truncated isoform of the thyrotropin-releasing hormone receptor is expressed in the rat central nervous system as well as in the pituitary gland.";
 RL Brain Res. Mol. Brain Res. 20:353-356(1993).
 RN 17
 RC TISSUE: FROM N.A.
 RX STRAIN=WISTAR;
 RX MEDLINE=92231953; PubMed=1373613;
 RA Yamada M., Monden T., Satoh T., Itzuka M., Murakami M., Iriuchijima T., Mori M.;
 RT "Differential regulation of thyrotropin-releasing hormone receptor mRNA levels by thyroid hormone in vivo and in vitro (GH3 cells).";
 RL Biochem. Biophys. Res. Commun. 184:367-372(1992).
 RN 18
 RP ALTERNATIVE SPLICING.
 RX MEDLINE=93100278; PubMed=1334485;
 RA de la Pena P., Delgado L.M., del Camino D., Barros F.;
 RT "Two isoforms of the thyrotropin-releasing hormone receptor generated by alternative splicing have indistinguishable functional properties.";
 RL J. Biol. Chem. 267:25703-25708(1992).
 CC -I- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS: TRH-R(412) (SHOWN HERE) AND TRH-R(387); ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M90308; AAA42277.1; -
 DR EMBL: X64630; CAA45913.1; -
 DR EMBL: X66726; CAA47263.1; -
 DR EMBL: D17469; BAA04289.1; -
 DR EMBL: S60053; AAB26491.1; -
 DR EMBL: S69160; AAB29945.2; -
 DR EMBL: S69161; AAB29946.2; -
 DR EMBL: S51512; AAB24549.1; -
 DR PIR: S23436; S23436.
 DR PIR: P00326; P00326.
 DR GCRDB: GCR_0306; -
 DR GCRDB: GCR_0431; -
 DR GCRDB: GCR_0691; -
 DR GCRDB: GCR_0943; -
 DR GCRDB: GCR_1031; -
 DR GCRDB: GCR_1299; -
 DR GCRDB: GCR_1310; -
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1
 DR PRINTS: PR00751; THYROLIBRIN.
 DR PROSITE: PS00237; G_PROTEIN_REC_F1_1; 1.
 DR PROSITE: PS50262; G_PROTEIN_REC_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Alternative splicing.

```

FT DOMAIN 1 28 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 29 51 1 (POTENTIAL).
FT DOMAIN 52 61 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 62 83 2 (POTENTIAL).
FT DOMAIN 84 93 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 100 121 3 (POTENTIAL).
FT DOMAIN 122 144 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 145 168 4 (POTENTIAL).
FT DOMAIN 169 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215 5 (POTENTIAL).
FT DOMAIN 216 266 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 267 288 6 (POTENTIAL).
FT DOMAIN 289 296 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 297 319 7 (POTENTIAL).
FT DOMAIN 320 412 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 376 412 YVSTKVSFDITCLASEKNGPSSCTGYSLTAKQKXI ->
RMDPVLVHMDIL (IN ISOFORM TRH-R(387)).
FT CONFLICT 59 59 A -> P (IN REF. 6 AND 7).
FT CONFLICT 223 223 I -> T (IN REF. 6 AND 7).
FT CONFLICT 321 321 N -> T (IN REF. 6).
FT CONFLICT 393 393 MISSING (IN REF. 3 AND 4).
SQ SEQUENCE 412 AA; 46608 MW; 52B35C9849D4B72 CRC64;

```

Query Match 18.0%; Score 393; DB 1; Length 412;
 Best Local Similarity 29.8%; Pred. No. 1,2e-17;
 Matches 111; Conservative 69; Mismatches 116; Indels 74; Gaps 13;

```

OY 18 LEDPFOKHLNSTEELYLAFLCGRSHFPLVSVVVPFV- -GVIGNVLCVLIQH 73
DB 1 MENETVSELNQTLEL- - - - -PQVAVALEYQVITLVVIGLGIVGINVAVLMRT 53
OY 74 QAKKPTNYLPSLANSDDLVL- -LCMP- - -EYEMKRYPLFEPGVCYFTALFEY 129
DB 54 KKHRTATNGYLVSLAVADLVLAAGLPNTDSIYGSW- - - - -VYGVGLCTITYLOYLG 108
OY 130 CPASLITTVSVERYVALHPRAKLOSTRRALRILGIVGFSVLSLPNTSINGIKF 189
DB 109 INASSCSITAFTERIALICHPKAOFLCTFSKAKIITFVAFTSYICMLNFFLDLNI 168
OY 190 HYFNGSLVPGSATC- - - - -TVIKPMIYVNFIIQVTSFELLPMTVISLVLYMALRL 243
DB 169 STYKDAIVI- - -SCGYKISRNYSPILMDFGV- - -FYVPMILATVLGFIARIL 218
OY 244 - - - - -KKKKSLEADGNNANIORC- - - - -KSNKMLFVLVFAIC 280
DB 219 FLNPIPSDKENSKYKNDSTHQNKNNMLTNTNRCFNSVSVSRKQYTKMLAVVILFALL 278
OY 281 MAPFH- - -IDRLFFS- - -FVEEMSESLAAVFNLVHVGVEFYLSAVNPITNLSRRFQ 335
DB 279 WMPYRRLVYVNSFLSPFOENW- - - - -FLFCRCICITLNSAINPIYTLMSQKFR 328
OY 336 AAFQNVISFHK 347
DB 329 AAFKRLCNCKOK 340

```

RESULT 15
 TRFR_BOVIN STANDARD; PRT; 398 AA.
 AC 046639;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE THYROTROPIN-RELEASING HORMONE RECEPTOR (TRH-R) (THYROLIBERIN RECEPTOR).
 GN TRHR.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.

```

OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HOLSTEIN;
RX MEDLINE=98151357; PubMed=9492373;
RA Takata M., Shimada Y., Ikeda A., Sekikawa K.;
RT "Molecular cloning of bovine thyrotropin-releasing hormone receptor gene."
RL J. Vet. Med. Sci. 60:123-127(1998).
CC -!- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC -----
DR EMBL; D83964; BAA24069.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00751; THYROLIBRINR.
DR PROSITE; PS00237; G_PROTEIN_RECPT_FL_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECPT_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 28 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 29 51 1 (POTENTIAL).
FT DOMAIN 52 61 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 62 83 2 (POTENTIAL).
FT DOMAIN 84 93 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 100 121 3 (POTENTIAL).
FT DOMAIN 122 144 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 145 168 4 (POTENTIAL).
FT DOMAIN 169 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215 5 (POTENTIAL).
FT DOMAIN 216 266 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 267 288 6 (POTENTIAL).
FT DOMAIN 289 296 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 297 319 7 (POTENTIAL).
FT DOMAIN 320 398 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 398 AA; 45165 MW; D9AF4B21A5701B8 CRC64;

```

Query Match 17.8%; Score 389.5; DB 1; Length 398;
 Best Local Similarity 29.5%; Pred. No. 1,9e-17;
 Matches 112; Conservative 73; Mismatches 112; Indels 83; Gaps 14;

```

OY 48 VSVVYVPIEV- -GVIGNVLCVLIQHQAOKTPTNYLPSLAVSDDLVL- -LCMP- - -E 102
DB 27 VTILVLIIGLGIVGINVAVLMVMTKHMRTFNCYLSLAVADLVLAAGLPNTDS 86
OY 103 VYEMKRYPLFEPGVCYFTALFEYVCFASLITTVSVERYVALHPRAKLOSTRR 162
DB 87 IYGSW- - - - -VYGVGLCTITYLOYLGINSVSCSITAFTERIALICHPKAOFLCTFSR 141
OY 163 ALRILGIVGFSVLSLPNTSINGIKFHYFNGSLVPGSATC- - - - -TVIKPMIYVNF 216
DB 142 AKKIIIFVMAFTSYICMLNFFLDLNI STYKDAIV- - -SCGYKISRNYSPILMDFG 197
OY 217 IQVTSFELVLPMTVISLVLYMALRL- - - - -KKKKSLEADGNN 255
DB 198 V- - - - -FYVPMILATVLGFIARILFLNPIPSDKENSKYKNDSTHQNKNNLSKTSN 251
OY 256 A- - -NIORPKSKSVNKMFLVVLVFAICMAPFH- - -IDRLFFS- - -FVEEMSESLAAVFN 308
DB 252 RYFNSTVSVSRKQYTKMLAVVILFALLMMPYRRLVYVNSFLSPFOENW- - - - -F 301

```

Oy 309 HVVSGVFYLSAVNPITYNLRRFOAFONVSSFHQWHSQHDPOLPPAQRNI----- 364
Db 302 LLECRICITYLNSAINPVYINLMSOKFRAAFKILCNCOK-----PVEKPANYVALSY 354
Oy 365 -----FLTECHEVELTE 376
Db 355 SVIKESDRFSTELDDVYTD 374

Search completed: April 22, 2002, 14:19:41
Job time: 144 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:12:47 ; Search time 25.8 Seconds
(without alignments)
1225.288 Million cell updates/sec

Title: US-09-609-146-4
Perfect score: 2185
Sequence: 1 MSGMERLQNSMRYQKLED.....ALSSQMSRTNYSFHNKT 415

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	489.5	22.4	424	2 JH0164	neurotensin recept
2	480	22.0	378	2 T15816	hypothetical prote
3	471	21.6	418	2 S29506	neurotensin recept
4	410	18.6	416	2 S68822	neurotensin recept
5	405.5	18.6	398	2 JN0708	thyrotropin-releas
6	397	18.2	393	2 A39251	thyrotropin-releas
7	393	18.0	411	2 I56444	thyrotropin-releas
8	393	18.0	412	2 S23436	thyrotropin-releas
9	372	17.0	402	2 I56595	thyrotropin recept
10	365	16.7	369	2 B41795	neurokinin 2 recept
11	359.5	16.5	369	2 JC2083	somatostatin recept
12	358	16.4	369	2 A45291	somatostatin recept
13	357.5	16.4	399	2 S29480	bombesin recept
14	357	16.3	392	2 S65693	opioid receptor mu
15	357	16.3	398	2 A57510	opioid receptor mu
16	356.5	16.3	352	2 I56296	thyrotropin releas
17	355	16.2	400	2 I56553	mu opiate receptor
18	352.5	16.1	380	2 S36143	kappa opioid recept
19	351.5	16.1	385	2 S55524	neurokinin 3 recept
20	351	16.1	384	2 A47249	brain-specific som
21	350.5	16.0	380	2 JC2338	kappa opioid recept
22	350.5	16.0	380	2 A55259	kappa opioid recept
23	350	16.0	384	2 JC4629	somatostatin recept
24	349	16.0	369	2 D41795	somatostatin recept
25	348.5	15.9	380	2 A48227	kappa opioid recept
26	348.5	15.9	380	2 A38271	serotonin recept
27	347.5	15.9	391	2 A41795	somatostatin recept
28	347.5	15.9	391	2 C41795	somatostatin recept
29	347.5	15.9	391	2 A39297	somatostatin recept

30	347	15.9	394	2 JC7209	galanin receptor -
31	346.5	15.9	363	2 I57940	somatostatin recept
32	346	15.8	326	2 S29248	somatostatin recept
33	345.5	15.8	452	2 A34916	neurokinin 3 recept
34	344.5	15.8	398	2 I56517	mu opioid receptor
35	343.5	15.7	367	2 I56520	g protein-coupled
36	343	15.7	384	2 S20303	neurokinin 2 recept
37	342.5	15.7	399	2 A46632	bombesin-like pept
38	341.5	15.6	380	2 JC2434	kappa opioid recept
39	341	15.6	388	2 JN0605	somatostatin recept
40	340.5	15.6	367	2 I49022	kappa opioid recept
41	340.5	15.6	367	2 JC2421	opioid receptor ho
42	339.5	15.5	384	2 S00516	neurokinin 2 recept
43	339.5	15.5	465	1 J01517	neurokinin 3 recept
44	339	15.5	398	1 J01059	neurokinin 2 recept
45	338	15.5	418	2 A46226	somatostatin recept

ALIGNMENTS

RESULT 1
JH0164
neurotensin receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 17-Mar-2000
C:Accession: JH0164
R:Tanaka, K.; Masu, M.; Nakanishi, S.
Neuron 4, 847-854, 1990
A:Title: Structure and functional expression of the cloned rat neurotensin receptor.
A:Reference number: JH0164; PMID:90297956
A:Accession: JH0164
A:Molecule type: mRNA
A:Residues: 1-424 <TRAN>
C:Comment: Neurotensin receptor belongs to the family of G protein-coupled receptor.
lter (neurotransmitter) in the brain and as a hormone) cellular mediator in peripheral t
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F:97-121/Domain: transmembrane #status predicted <TM2>
F:144-165/Domain: transmembrane #status predicted <TM3>
F:189-210/Domain: transmembrane #status predicted <TM4>
F:236-260/Domain: transmembrane #status predicted <TM5>
F:309-330/Domain: transmembrane #status predicted <TM6>
F:348-372/Domain: transmembrane #status predicted <TM7>
F:4,38,42/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	22.4%	Score 489.5;	DB 2;	Length 424;
Best Local Similarity	32.3%	Pred. No. 5,1e-34;		
Matches 107;	Conservative 75;	Mismatches 114;	Indels 35;	Gaps 9;
OY	48	VSVVYVIEFVGVYIGNVLCVLIQH--QAKPTNYVLFSLAVSDLLVLLGMPLEVV	104	
DB	67	VTAIVYLFVGVGVGVSVTAFTLARKKSLQSLSTVHYHGLSLDILLAMPVELY	126	
OY	105	E-MWRNVPFLGPGVC--YFKTALFETVCFASLTSTTVSVEVYVLLIPPAKLOSTR	160	
DB	127	NFTVYHNPWFAGAGCGYTF--LRDCTYATALNANSLVEVYLLICHPFAKTLMSR	183	
OY	161	RRALRIIGVGFVSFLSPNTSINGIKFYFPGNSLVPSANCTVYKPMIYFIQVY	220	
DB	184	SRKKKFSALIMLSALLAIFMLFTMGLONR--SGDGNHPGIVCTPVDVATGKVIQVY	241	
OY	221	SLFLYLLPMTVIVSVLYLMLARKKDSLEADG-----NANTQPCRK	264	
DB	242	TFSFLEFPMIVISLINTVINANKLTVMVHQAAEDGRCVCTVGNGLHSTFNMTIE-GRV	300	
OY	265	SVVK---MLFVIVIVAIQMARPHIDRLFFESV--FENSESLAAYNVLHVVSQVFFYL	318	
DB	301	QALRHGVILRAVIVIAVAVVCMPLRHVRIMFYISDQMTTFLEDFVHYPMILNALLEVY	360	
OY	319	SSAVNPITVYMLSRFQAFQNVISFHKQW	349	

Db 361 SSAINPILYNLVSANFRQVFLSTLACCPGW 391

RESULT 2

T15816 hypothetical protein C48C5.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000

C:Accession: T15816

R:Ravello, A.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid C48C5.

A:Reference number: Z18410

A:Accession: T15816

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-378 <P>

A:Cross-references: EMBL:U039994; NID:91055102; PID:91055105; PIDN:AAB37017.1; GSPDB:GNOC

A:Experimental source: strain Bristol N2; clone C48C5

C:Genetics:

A:Gene: CESP-C48C5.1

A:Map position: X

A:Insertions: 53/3; 87/2; 136/2; 169/3; 209/3; 231/3; 259/3; 286/1; 327/1

C:Superfamily: adenosine receptor A1

Query Match 22.0%; Score 480; DB 2; Length 378;

Best Local Similarity 31.8%; Pred. No. 2.9e-33;

Matches 118; Conservative 67; Mismatches 146; Indels 40; Gaps 8;

QY 23 QKHLNSTEER---LAFICGPRRS-----FLLPVSVVYPIFV 56
 Db 3 QALNTEDDCCDLAFNCPIYVSHSESEKACVMEHCFISKRALDDVTLKVTALYLFIF 62
 QY 57 VGVGVNVLVCLVYLIOHQAOKPTNYLFSGLAVSDLLVLLGMPLEVEYEMRNYPFLFGP 116
 Db 63 LVGVIGNTTCLVWKKRPMKTHASMYLMLAVSDVLTLCVGLPEVYMMNMNQYPMFPDP 122
 QY 117 VGVYFKTALFETVCFASISITVSVERYVAIIHP--FRALQSTRRALILIGVWGFVS 175
 Db 123 YICNLKALMETSSVILITILFAIRRYAVACHPLFMVOPKRNIGITIGTWTFFSI 182
 QY 176 LFLSLPNTSINGIKF--HYFP---NGSLVPSANCTVI-----KPMWYNYFIQVTSLEY 225
 Db 183 LCAFPFIIHRADYIMKSMFGTDNRIPVKSCKMIMAVMEPKIASFFKILFHSAIAFF 242
 QY 226 LHPMTVSVLYLMLALRLKDKSLDEAGNANIORPCRSVKNMLFVLVLFVAFICWAPFH 285
 Db 243 ALPLFTIVILYARIACVSSNRTIQ--PGELDITEELQMRIMAILCAIVSAFFICYLPFQ 300
 QY 286 IDPLFSEFSEWSESLAAVENLVHVSQVFFYSSAVNPITNYNLSRPOAQRONVSSSF 345
 Db 301 LQRLLPFFYFDN-EVILTWVQMYMIFISGLFLYATITINPIAVNLASSRFRRAFKDILIDY 359
 QY 346 HKQWHSQHPDQ 356
 Db 360 CMRGSSERYPR 370

RESULT 3

S29506 neurotensin receptor - human

C:Species: Homo sapiens (man)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-2000

C:Accession: S29506

R:Vita, N.; Laurent, P.; Lefort, S.; Chalou, P.; Dumont, X.; Kaghad, M.; Gully, D.; Le F

EBES Letc. 317, 139-142, 1993

A:Title: Cloning and expression of a complementary DNA encoding a high affinity human ne

A:Reference number: S29506; MUID:93154505

A:Accession: S29506

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-418 <VT>
 A:Cross-references: EMBL:X70070; NID:935020; PIDN:CAA9675.1; PID:935021
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 21.6%; Score 471; DB 2; Length 418;

Best Local Similarity 31.0%; Pred. No. 1.9e-32;

Matches 120; Conservative 72; Mismatches 121; Indels 74; Gaps 14;

QY 20 DPFQKHLNSTEYL-----AFICGPRRS-----HFLPVSVVYPIFV 58
 Db 17 DPFQAGLELEALDAGFCNAGNASERVLAAPSSSELDVDTQYSLVLTAVYALFV 76
 QY 59 GVIGNVLVCLVYLIOH---QAMKTPTNYLFSGLAVSDLLVLLGMPLEVEYEMRNYPFLF 114
 Db 77 GYGNVTVAFTLARKKSLQSLQSTVYHLSLALSDLTLLAMPVLELYNEIWHNMAF 136
 QY 115 GPVGC---YFKTALFETVCFASISITVSVERYVAIIHPFRALQSTRRALILIGVW 171
 Db 137 GDACRCGYTF---LRDCTYATALNVASLSYERLAIChFPKATLMSRSRTKKFISAIV 193
 QY 172 GFSVLFSLP-----NTSINGIKFHYFPNGSLV--PGSATCTVIKPMWYNYFIQVTS 221
 Db 194 LASALLTPVPLMTWGEQNRSDQCH---AGLVCPTIHTATV-----KVVIQVNT 241
 QY 222 FLPLLPMTVSVLYLMLALRLKKDKSLDEAGN-----ANQRCRK 264
 Db 242 FMSITFPVAVISVNTIANKLTVVVRQAEQGVCTVGEHSTFSAIEGRQV-ALRH 300
 QY 265 SVNKLFLVLVFAICWAPFHIDRLFFSFV--EEMSESLAAVENLVHVSQVFFYSSAV 322
 Db 301 GV-RVLAAYVIAFVYVCLPPIHVRRLMFCYISDEGMTFPLVDYHYFYVYNALFYVSSI 359
 QY 323 NPITYNLSRFPQAFQNVISSEHKQW 349
 Db 360 NPILYNLVSANFRHIFLATLACCPW 386

RESULT 4

S68822 neurotensin receptor 2, leucobastine-sensitive - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 20-Jun-2000

C:Accession: S68822

R:Chalon, P.; Vita, N.; Kaghad, M.; Gullémot, M.; Bonnin, J.; Delpech, B.; Le Fur, G

EBES Letc. 386, 91-94, 1996

A:Title: Molecular cloning of a leucobastine-sensitive neurotensin binding site.

A:Reference number: S68822; MUID:96228041

A:Accession: S68822

A:Molecule type: mRNA

A:Residues: 1-416 <CHA>

A:Cross-references: GB:X97121; NID:91483579; PIDN:CAM65787.1; PID:91483580

A:Experimental source: hypothalamus

C:Superfamily: vertebrate rhodopsin

C:Keywords: G protein-coupled receptor; transmembrane protein

F:33-58/Domain: transmembrane #status predicted <TM1>

F:70-91/Domain: transmembrane #status predicted <TM2>

F:110-131/Domain: transmembrane #status predicted <TM3>

F:155-175/Domain: transmembrane #status predicted <TM4>

F:204-230/Domain: transmembrane #status predicted <TM5>

F:296-315/Domain: transmembrane #status predicted <TM6>

F:335-361/Domain: transmembrane #status predicted <TM7>

Query Match 18.8%; Score 410; DB 2; Length 416;

Best Local Similarity 30.5%; Pred. No. 2.8e-27;

Matches 107; Conservative 67; Mismatches 125; Indels 52; Gaps 10;

QY 49 SVYVYVFFVGVGNVLVCLVYLIOHQAOKPTNYLFSGLAVSDLLVLLGMPLEVEYEM 106
 Db 36 TALYSILFAFGTAGNALSVHVLKARAGRGRLRYHVLSTLALSLLLVSMPELYNFV 95

OY 107 WRNYPFLFGPVGC---YFKTALFETVCFASILSTITVSEVERVAAILHPFRAKLQSTRRRA 163
 F:101-121/Domain: transmembrane #status predicted <TM2>
 E:146-168/Domain: transmembrane #status predicted <TM3>
 F:194-215/Domain: transmembrane #status predicted <TM4>
 F:267-288/Domain: transmembrane #status predicted <TM5>
 F:297-319/Domain: transmembrane #status predicted <TM6>
 Db 153 RRLSLVWVASLGLALPMVAVIMQKHEVESADEPEPASVCVIVLSRATLQVFIQVNVL 212
 OY 223 LFYLPMTVTSVLSV-----YYLMLA-----RLKKDKSLDEADG----- 254
 Best Local Similarity 29.9%; Pred. No. 6, 5e-27; Length 398;
 Matches 119; Conservative 72; Mismatches 130; Indels 77; Gaps 14;
 Db 213 VSFALPLALFALFNGITVNHMLVLSQVPSASAOVSSISRLLELSEEGGLGITWRKTL 272
 OY 255 -----NANTQPCRSKVNKMLFVLVFAICWAPFHIDRLFSFVEE--WSES 300
 F:146-168/Domain: transmembrane #status predicted <TM3>
 F:194-215/Domain: transmembrane #status predicted <TM4>
 F:267-288/Domain: transmembrane #status predicted <TM5>
 F:297-319/Domain: transmembrane #status predicted <TM6>
 Db 273 SLGVQASLVRHKDAQISRLQSHA-QVLRALVAVVYICMLPYHARLRMLCYIPDDQWTNE 331
 OY 301 LAAVFLVAVVSGVFYSSAVNPITLYNLRRFOAFQNVISSEFKQWHS 351
 F:101-121/Domain: transmembrane #status predicted <TM2>
 E:146-168/Domain: transmembrane #status predicted <TM3>
 F:194-215/Domain: transmembrane #status predicted <TM4>
 F:267-288/Domain: transmembrane #status predicted <TM5>
 F:297-319/Domain: transmembrane #status predicted <TM6>
 Db 332 LVDFVHYFVWVNTLFYVSSAVTPILYNAVSSSFRLFLFSLGSLGEGHS 382
 RESULT 5
 JN0708
 thyrotropin-releasing hormone receptor - human
 C:Species: Homo sapiens (man)
 C:Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Jun-2000
 C:Accession: S40682; JN0759; S50151; S50152; I38356; JN0708
 R:Matte, V.; Katsen, H.E.; Wright, M.S.; Lundell, I.; Fjeldheim, A.K.; Gabrielsen, O.S.
 Blochm. Biophys. Res. Commun. 195, 179-185, 1993
 A:Title: Molecular cloning of a functional human thyrotropin-releasing hormone receptor.
 A:Reference number: S40682; MUID:93371401
 A:Accession: S40682
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-398 <MAT>
 A:Cross-references: EMBL:X75071; NID:9404157; PIDN:CAA52965.1; PID:9404158
 R:Yamada, M.; Monden, T.; Satoh, T.; Satoh, N.; Murakami, M.; Iriuchijima, T.; Kakegawa,
 Blochm. Biophys. Res. Commun. 195, 737-745, 1993
 A:Title: Pituitary adenomas of patients with acromegaly express thyrotropin-releasing ho
 rone gene.
 A:Reference number: JN0759; MUID:93384596
 A:Accession: JN0759
 A:Molecule type: mRNA
 A:Residues: 1-398 <YAM>
 A:Cross-references: GB:D16845; NID:9577631; PIDN:BA04120.1; PID:9577632
 R:Hluma, S.; Hosoya, M.; Ogi, K.; Tanaka, H.; Negat, Y.; Onda, H.
 Blochm. Biophys. Res. Commun. 195, 251-255, 1994
 A:Title: Molecular cloning and functional expression of a human thyrotropin-releasing ho
 rone gene.
 A:Reference number: S50151; MUID:95002135
 A:Accession: S50151
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-263 <HIN>
 A:Accession: S50152
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 267-398 <HI2>
 R:Dutrie, S.M.; Taylor, P.L.; Anderson, L.; Cook, J.; Eldine, K.A.
 Mol. Cell. Endocrinol. 95, R11-R15, 1993
 A:Title: Cloning and functional characterisation of the human TRH receptor.
 A:Reference number: I38356; MUID:94063224
 A:Accession: I38356
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-398 <RES>
 A:Cross-references: EMBL:X72089; NID:9440155; PIDN:CAA50979.1; PID:9440156
 C:Genetics:
 A:Gene: GDB:TRHR
 A:Cross-references: GDB:228955; OMIM:188545
 A:Map position: 8q23-8q23
 C:Superfamily: adenosine receptor A1
 C:Keywords: G protein-coupled receptor; receptor; transmembrane protein
 F:29-51/Domain: transmembrane #status predicted <TM1>

F:62-83/Domain: transmembrane #status predicted <TM2>
 F:101-121/Domain: transmembrane #status predicted <TM3>
 F:146-168/Domain: transmembrane #status predicted <TM4>
 F:194-215/Domain: transmembrane #status predicted <TM5>
 F:267-288/Domain: transmembrane #status predicted <TM6>
 F:297-319/Domain: transmembrane #status predicted <TM7>
 Query Match 18.6%; Score 405.5; DB 2; Length 398;
 Best Local Similarity 29.9%; Pred. No. 6, 5e-27; Length 398;
 Matches 119; Conservative 72; Mismatches 130; Indels 77; Gaps 14;
 OY 48 VSVVYVPIEV-VGVICNVLCVILQHOAKPTPNYYLSLAASDLVLVL-LGMPL---E 102
 F:146-168/Domain: transmembrane #status predicted <TM3>
 F:194-215/Domain: transmembrane #status predicted <TM4>
 F:267-288/Domain: transmembrane #status predicted <TM5>
 F:297-319/Domain: transmembrane #status predicted <TM6>
 Db 27 VTIILVLIICGLGIVGINVAVLVMTKHKRTPTNCLVSLAVADLMLVLAAGLPNTDS 86
 OY 103 VEDMKRNPFLFGPVGCYFETVCFASILSTITVSEVERVAAILHPFRAKLQSTRR 162
 F:101-121/Domain: transmembrane #status predicted <TM2>
 E:146-168/Domain: transmembrane #status predicted <TM3>
 F:194-215/Domain: transmembrane #status predicted <TM4>
 F:267-288/Domain: transmembrane #status predicted <TM5>
 F:297-319/Domain: transmembrane #status predicted <TM6>
 Db 87 IYGSW-----VGVYVCLCTIYQLQGINASSCSITAFTERIYAIICHPIKAQFLCTFSR 141
 OY 163 ALRILGIVGFSVLSLPNTSHIGIKFHPNSLYPGSATC-----IYIKPMIYNI 216
 F:146-168/Domain: transmembrane #status predicted <TM3>
 F:194-215/Domain: transmembrane #status predicted <TM4>
 F:267-288/Domain: transmembrane #status predicted <TM5>
 F:297-319/Domain: transmembrane #status predicted <TM6>
 Db 142 AKRIIFVMAFSLVLCMLFELDLNISTYKDAIV---SCGYKISRNYSPIYLMDDG 197
 OY 217 IQVTSFELFLPMTVTSVLYLMLRL-----KDKSLDEADGNANIOR 260
 F:146-168/Domain: transmembrane #status predicted <TM3>
 F:194-215/Domain: transmembrane #status predicted <TM4>
 F:267-288/Domain: transmembrane #status predicted <TM5>
 F:297-319/Domain: transmembrane #status predicted <TM6>
 Db 198 V-----FVYVPIILATVILGFIARLLFLNPISDPKRENSKTKNDSTHONTNLTNS 251
 OY 261 PC-----RKSVMKLFVLVFAICWAPFH--IDRLFFS--EVEENSESLAAVFNLY 308
 F:146-168/Domain: transmembrane #status predicted <TM3>
 F:194-215/Domain: transmembrane #status predicted <TM4>
 F:267-288/Domain: transmembrane #status predicted <TM5>
 F:297-319/Domain: transmembrane #status predicted <TM6>
 Db 252 RCENSTVSSRKQVTKMLAVVILFALLMPYRFLVAVNSPLSPQENW-----F 301
 OY 309 HAVSGVFYLLSAVNPITLYNLRRFOAFQNVISSEFKQWHSQHOPOLPRAORNIPL 367
 F:146-168/Domain: transmembrane #status predicted <TM3>
 F:194-215/Domain: transmembrane #status predicted <TM4>
 F:267-288/Domain: transmembrane #status predicted <TM5>
 F:297-319/Domain: transmembrane #status predicted <TM6>
 Db 302 LFCRCITVLSAIPVITVILMSQKFRFAERKICNCKOK-----PTKPNVYSLAVNY 354
 OY 368 -----ECHFEVELTEDEGPQPCQSMHNSHLPALSSSE 400
 F:146-168/Domain: transmembrane #status predicted <TM3>
 F:194-215/Domain: transmembrane #status predicted <TM4>
 F:267-288/Domain: transmembrane #status predicted <TM5>
 F:297-319/Domain: transmembrane #status predicted <TM6>
 Db 355 SVIKESDHPSTELDDITVDITVTSATKVSFDDCLASE 392
 RESULT 6
 A39251
 thyrotropin-releasing hormone receptor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 11-Jan-2000
 C:Accession: A39251
 R:Straub, R.E.; Frech, G.C.; Joho, R.H.; Gershengorn, M.C.
 Proc. Natl. Acad. Sci. U.S.A. 87, 9514-9518, 1990
 A:Title: Expression cloning of a cDNA encoding the mouse pituitary thyrotropin-releas
 ing hormone receptor.
 A:Reference number: A39251; MUID:91088548
 A:Accession: A39251
 A:Molecule type: mRNA
 A:Residues: 1-393 <STR>
 A:Cross-references: GB:M59811; GB:M37490; NID:9202153; PIDN:AAA40480.1; PID:9202154
 C:Superfamily: adenosine receptor A1
 C:Keywords: G protein-coupled receptor; glycoprotein; membrane protein
 Query Match 18.2%; Score 397; DB 2; Length 393;
 Best Local Similarity 28.6%; Pred. No. 3, 4e-26;
 Matches 110; Conservative 72; Mismatches 104; Indels 98; Gaps 15;
 OY 18 LEDPFGKHLNSTE-----EY-----LAFLCGPRRSHFELVSVVYVPIFVGYI 61
 F:146-168/Domain: transmembrane #status predicted <TM3>
 F:194-215/Domain: transmembrane #status predicted <TM4>
 F:267-288/Domain: transmembrane #status predicted <TM5>
 F:297-319/Domain: transmembrane #status predicted <TM6>
 Db 1 MENDTVSEKMQTLEDOAAVALEXYQVVTLLVLIIG-----LGIV 41
 OY 62 GNVLCVLIILQHOAKPTPNYYLSLAASDLVLVL-LGMPL---EYEMKRNYPFLFGPV 117
 F:146-168/Domain: transmembrane #status predicted <TM3>
 F:194-215/Domain: transmembrane #status predicted <TM4>
 F:267-288/Domain: transmembrane #status predicted <TM5>
 F:297-319/Domain: transmembrane #status predicted <TM6>
 Db 42 GNTMVVAVNKRTHKRTPTNCLVSLAVADLMLVLAAGLPNTDSYGSN-----YGVY 96
 OY 118 GCYFKTALFETVCFASILSTITVSEVERVAAILHPFRAKLQSTRRRLILGIVGFSVLP 177
 F:101-121/Domain: transmembrane #status predicted <TM2>
 E:146-168/Domain: transmembrane #status predicted <TM3>
 F:194-215/Domain: transmembrane #status predicted <TM4>
 F:267-288/Domain: transmembrane #status predicted <TM5>
 F:297-319/Domain: transmembrane #status predicted <TM6>

Db 97 GCICITVLYOYLGINASSCSITAFIERIYAICHPIKAQFLCTFSRAKIIIFVMAFTSIY 156
 QY 178 SLPTNSIHGKIFHYFPGNSIVPGSATC-----TVIKPMIYVFIQVTSFLVLLPMV 231
 Db 157 CMKMFLLDLNISTYKNAVY-----SCGYKISRNYSPILMDFGV-----FYVPMIL 206
 QY 232 ISVLYLYMA-----LRLKDKSLEADG--NANIQRPC-----RKSVMK 268
 Db 207 ATVLYGFIARILFLNPISDPKNSKMKMKNDISHONKNLNLNPTNCFNSTVSSRKQVTK 266
 QY 269 MFLVLLVFAICAPRH---IDRLFFS--FVEKSESLLAANLVHVSCEFYLSSAVN 323
 Db 267 MAAVYVILFLFLMMPYRTLVVNSFLSPQENW-----FLLCRICIYLNSAIN 316
 QY 324 PIYVNLSSRFOAFONVSSFK 347
 Db 317 PVLYNLSQKFRFAFKLNCCKK 340

RESULT 7
 I56444
 thyrotrophin-releasing hormone receptor - mouse
 C:Species: Mus sp. (mouse)
 C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 11-Jan-2000
 C:Accession: I56444
 R:SELLER, R.E.; TAYLOR, P.L.; LAMB, R.F.; ZABAVNIK, J.; ANDERSON, L.; ELDNE, K.A.
 J. Mol. Endocrinol. 10, 199-206, 1993
 A:Title: Functional expression and molecular characterization of the thyrotrophin-releasing hormone receptor
 A:Reference number: I56444; MIMD:93249585
 A:Accession: I56444
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-411 <RES>
 A:Cross-references: S60053; NID:9300151; PIDN:AAB26491.1; PID:9300152
 C:Superfamily: adenosine receptor A1

Query Match 18.0%; Score 393; DB 2; Length 411;
 Best local similarity 29.8%; Pred. No. 7.8e-26;
 Matches 111; Conservative 69; Mismatches 118; Indels 74; Gaps 13;

QY 18 LEDPFOKHLNSTEYIAFLCGPRRSHEFLVSVYVPIFYV---GVIGNVLCVLILQH 73
 Db 1 MENETVSELNQTEL-----PQVAVALEYQVVTILLVVICGLIGVINIMVLLVMRT 53
 QY 74 QAKTPPNVYLFSLAVSLLVLL-LGMPL---EYEMKRWYPLFLGVCYCFYTALEFV 129
 Db 54 KHMRTATNCLVSLAVADLVLAAGLNPITDSIYGSW-----VYGVGCLCTIYLOYL 108
 QY 130 CPASIIITTVSVERYVAIILHPRAKLOSTRRALRIILGIWGSVFLSPNTSIHGKIF 189
 Db 109 INSSCSITAFTEIRYIAICHPIKAQFLCTFSRAKIIIFVMAFTSIYCMLEPFLDLNI 168
 QY 190 HYFPNGSLVGSATC-----TVIKPMIYVFIQVTSFLVLLPMVTSVLYLYMALRL 243
 Db 169 STYKDAIV-----SCGYKISRNYSPILMDFGV-----FYVPMILATLVLGFIARIL 218
 QY 244 -----KKDSLEADGNANIQRPC-----RKSVMKLFVLVFAIC 280
 Db 219 FLNPISDPKNSKTKWKNNDSTHONKNMLNTNCFNSTVSSRKQVTKMLAVVILFALL 278
 QY 281 WAFPH---IDRLFFS--FVEKSESLLAANLVHVSCEFYLSSAVNPIIYNLSSRFQ 335
 Db 279 WMPYRLLVAVNSFLSPQENW-----FLLCRICIYLNSAINPVYINLMSOKFR 328
 QY 336 AAFONVSSFK 347
 Db 329 AAFRLKNCCKK 340

RESULT 8
 S23436
 thyroliberin receptor - rat

N:Alternate names: thyrotrophin-releasing hormone receptor
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 11-Jan-2000
 C:Accession: S23436; I53279; A49168; P00326
 R:de la Pena, P.; Delgado, L.M.; del Camillo, D.; Barros, F.
 Biochem. J. 284, 891-899, 1992
 A:Title: Cloning and expression of the thyrotrophin-releasing hormone receptor from GH
 A:Reference number: S23436; MIMD:9232017
 A:Accession: S23436
 A:Molecule type: mRNA
 A:Residues: 1-412 <PEN>
 A:Cross-references: EMBL:564630; NID:957394; PIDN:CAA45913.1; PID:957395
 R:Kimura, N.; Arai, K.; Sahara, Y.; Suzuki, H.; Kimura, N.
 Endocrinology 134, 432-440, 1994
 A:Title: Estradiol transcriptionally and posttranscriptionally up-regulates thyrotropin
 A:Reference number: I53279; MIMD:94102223
 A:Accession: I53279
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-412 <RES>
 A:Cross-references: GB:D17469; NID:9464199; PIDN:BA04289.1; PID:9464200
 R:Zhao, D.; Yang, J.; Jones, K.E.; Gerald, C.; Suzuki, Y.; Hogan, P.G.; Chin, W.W.; T
 Endocrinology 130, 3529-3536, 1992
 A:Title: Molecular cloning of a complementary deoxyribonucleic acid encoding the thy
 A:Reference number: A49168; MIMD:92283212
 A:Accession: A49168
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-12, 'D', '14-290', 'K', '292-412' <ZHA>
 A:Experimental source: GH cells
 A:Note: sequence extracted from NCBI backbone (NCBIN:104788, NCBI:104795)
 R:Yamada, M.; Monden, T.; Satoh, T.; Iizuka, M.; Murakami, M.; Iriuchijima, T.; Mori,
 Biochem. Biophys. Res. Commun. 184, 367-372, 1992
 A:Title: Differential regulation of thyrotrophin-releasing hormone receptor mRNA level
 A:Reference number: P00326; MIMD:92231953
 A:Accession: P00326
 A:Molecule type: mRNA
 A:Residues: 30-58, 'P', '60-222', 'T', '224-261' <YAM>
 A:Experimental source: strain Myster
 A:Note: the authors translated the codon ACA for residue 88 as Ala
 C:Superfamily: adenosine receptor A1
 C:Keywords: G protein-coupled receptor; transmembrane protein
 F:1-12/Domain: transmembrane #status predicted <TM1>
 F:32-54/Domain: transmembrane #status predicted <TM2>
 F:71-92/Domain: transmembrane #status predicted <TM3>
 F:116-140/Domain: transmembrane #status predicted <TM4>
 F:165-186/Domain: transmembrane #status predicted <TM5>

Query Match 18.0%; Score 393; DB 2; Length 412;
 Best local similarity 29.8%; Pred. No. 7.8e-26;
 Matches 111; Conservative 69; Mismatches 118; Indels 74; Gaps 13;

QY 18 LEDPFOKHLNSTEYIAFLCGPRRSHEFLVSVYVPIFYV---GVIGNVLCVLILQH 73
 Db 1 MENETVSELNQTEL-----PQVAVALEYQVVTILLVVICGLIGVINIMVLLVMRT 53
 QY 74 QAKTPPNVYLFSLAVSLLVLL-LGMPL---EYEMKRWYPLFLGVCYCFYTALEFV 129
 Db 54 KHMRTATNCLVSLAVADLVLAAGLNPITDSIYGSW-----VYGVGCLCTIYLOYL 108
 QY 130 CPASIIITTVSVERYVAIILHPRAKLOSTRRALRIILGIWGSVFLSPNTSIHGKIF 189
 Db 109 INSSCSITAFTEIRYIAICHPIKAQFLCTFSRAKIIIFVMAFTSIYCMLEPFLDLNI 168
 QY 190 HYFPNGSLVGSATC-----TVIKPMIYVFIQVTSFLVLLPMVTSVLYLYMALRL 243
 Db 169 STYKDAIV-----SCGYKISRNYSPILMDFGV-----FYVPMILATLVLGFIARIL 218
 QY 244 -----KKDSLEADGNANIQRPC-----RKSVMKLFVLVFAIC 280
 Db 219 FLNPISDPKNSKTKWKNNDSTHONKNMLNTNCFNSTVSSRKQVTKMLAVVILFALL 278

F:254-286/Domain: transmembrane #status predicted <TM6>
 F:292-316/Domain: transmembrane #status predicted <TM7>
 F:9,22,29,32/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:115-193/Disulfide bonds: #status predicted
 F:244,343/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinase II) #
 F:250/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predicted
 F:328/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 16.5%; Score 359.5; DB 2; Length 369;
 Best Local Similarity 29.3%; Pred. No. 4,9e-23;
 Matches 103; Conservative 67; Mismatches 138; Indels 43; Gaps 12;

QY 6 KLONASWYQKLEDPFO-----KHLNSTEYLAFLCGRSRHFLPVSVYVPIFV 57
 Db 6 ELNNGS---QPMLSFPDLNGSVATANSSNOTEPYDITSN-----AVLTFFYFVCI 55
 QY 58 VGVIGVNVCLVILQHQAMKTPNNYLFSLAVSDLLVLLGMP-LEVEEMRNYPFLGCP 116
 Db 56 IGLGNTLVYVILRYKMKMTINITYILNLADEL-FMLGLPFLAMQVALVHWP-FGK 112
 QY 117 VGGYFTALFEYVCFASLITTVSVRYVAIILHFRAKQSTRRALRLIGIYMGSVL 176
 Db 113 AICRVVMTVDGINQFTSIFCLTVMSIDRYLAVVHPKSAKRRPRAKMINVAVGVSLL 172
 QY 177 FSLPNTSIHGKHFHPNGSLVPSATCTVIKPR---MWIYNFIQVTSFELYLLPMTVI 232
 Db 173 VILPIMIVAGLRNQM-----GRSCTINMPGSGAWTGTFTI-YAFLGLFLVPLTII 224
 QY 233 SVLYYLMALRLKDKSLDEAGNANIQRPCRKYNMKLFVLVFAICMAPFHIDL-FF 291
 Db 225 CLCYELRIIVKVS---SGIRVGSKRKKSEKVTYRMVSIIVAVFICWLPFIYFNVS 280
 QY 292 SFVEEMSESLAAVENLVHVGVEFYLSAVNPPIYVNLSSRRQAFOAVI 342
 Db 281 SVAISFPALKGMDFEYVLT-----YANSCANPLIALFLSDNKKSFQNVL 327

RESULT 12

A45291
 somatostatin receptor, somatotropin release-inhibiting factor receptor, SKIF receptor -
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 25-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
 C:Accession: A45291
 R:Kluxen, F.W.; Bruns, C.; Lubbert, H.
 Proc. Natl. Acad. Sci. U.S.A. 89, 4618-4622, 1992
 A:Title: Expression cloning of a rat brain somatostatin receptor cDNA.
 A:Reference number: A45291; MUID:92262491
 A:Accession: A45291
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-369 <LNU>
 A:Cross-references: GB:M93773; NID:q207026; PIDN:AAA42165.1; PID:q2707027
 C:Note: sequence extracted from NCBI backbone (NCBIN:102315, NCBIPI:102316)
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 16.4%; Score 358; DB 2; Length 369;
 Best Local Similarity 30.1%; Pred. No. 6.5e-23;
 Matches 97; Conservative 65; Mismatches 128; Indels 32; Gaps 10;

QY 27 NSTEEXLAFLCGRSRHFLPVSVVYVPIFVGVGNVCLVILQHQAMKTPNNYLF 86
 Db 32 NQTEPYDITSN-----AVLTFFYFVGVGLGNTLVYILRLAKMKTTINITYILN 84
 QY 87 LAVSLLVLLGMP-LEVEEMRNYPFLGCPVCCYFETALFEYVCFASLITTVSVRY 145
 Db 85 LAIADEL-FMLGLPFLAMQVALVHWP-FGKALCRVVMVDGINQFTSIFCLTVMSIDRY 141
 QY 146 VALHFRAKLQSTRRALRLIGIYMGFSVLFSLPNTSIHGKHFHPNLSLVPGSATCT 205
 Db 142 LAVVHPKSAKRRPRAKMINVAVGVSLLVILPIMIVAGLRNQM-----GRSSCT 194

QY 206 VIKP-----MWIYNFIQVTSFELYLLPMTVSVLYLMALRLKDKSLDEAGNANIQR 261
 Db 195 INNPDESGAWTGTFTI-YAFLGLFLVPLTIIICLYLFIIVKVS---SGIRVGSKRKK 249
 QY 262 CRKSVNKMFLVYLVAICAPFHIDL-FFSFVEEMSESLAAVENLVHVGVEFYLS 320
 Db 250 SEKKVTRMVSIVAVFIFCLWLPFIYFNVSVAISPTPALKGMDFEYVLT-----YANS 305
 QY 321 AVNPIYNLSSRRQAFOAVI 342
 Db 306 CANPILYAFLLSDNKKSFQNVL 327

RESULT 13

S29480
 bombesin receptor - guinea pig
 C:Species: Cavia porcellus (guinea pig)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
 C:Accession: S29480
 R:Gorbulev, V.; Akhundova, A.; Buechner, H.; Fahrenholz, F.
 Submitted to the EMBL Data Library, July 1992
 A:Description: Molecular cloning of a new bombesin receptor subtype.
 A:Reference number: S29480
 A:Accession: S29480
 A:Molecule type: mRNA
 A:Status: preliminary
 A:Residues: 1-399 <GOR>
 A:Cross-references: EMBL:X67126; NID:g49545; PIDN:CAA47605.1; PID:g49546
 C:Superfamily: endothelin receptor B
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 16.4%; Score 357.5; DB 2; Length 399;
 Best Local Similarity 28.5%; Pred. No. 7.8e-23;
 Matches 85; Conservative 68; Mismatches 134; Indels 11; Gaps 4;

QY 48 VSVVYVPIFVGVGNVCLVILQHQAMKTPNNYLFSLAVSDLLVLLGMPLE-----V 103
 Db 50 IYITVAVIISVGLGNAILKVEFKTSMQTVNRIFITSLALGDLILLVCVVDATHYL 109
 QY 104 YEMWRNYPFLGPGCYEKTALFEYVCFASLITTVSVRYVAIILHFRAKLQSTRRA 163
 Db 110 AEGW-----LFGRIKCVLSFIRLTSGVSVFTLTISADRYKAVVYPLERQPSNALKT 164
 QY 164 LRLIGIYMGFSVLFSLPNTSIHGKHFHPNGSLVPSATCTVIKPMIYNFIQVTSFL 223
 Db 165 CANAGCIVMSMIFALPEALFSNVHTLRDPNKNMTSEMACFYPVSEKLOEIHALLSEFLV 224
 QY 224 FYLLPMTVSVLYLMLRL-KDKDSLDE-NNANIQRPCRKYNMKLFVLVFAICW 281
 Db 225 FYIIPLSIISVYSLIARTLYKSTLNIPTEQSHARKQVSRKRRIKTYLVLAALFALCW 284
 QY 282 APFHIDLRFSEVEMSESLAAVENLVHVGVEFYLSAVNPPIYVNLSSRRQAFO 339
 Db 285 LPMHLNLTLYHSFTYKAYEDSSAIFHTVITFSRYLAFNSCVNPFALYWLSTKTEQOKFK 342

RESULT 14

S65693
 opioid receptor mu variant MOR1A - human
 C:Species: Homo sapiens (man)
 C:Date: 12-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 05-Nov-1999
 C:Accession: S65693; S51216
 R:Bare, L.A.; Mansson, E.; Yang, D.
 Submitted to the EMBL Data Library, July 1994
 A:Description: Expression of two variants of the human mu opioid receptor mRNA in SK-
 A:Reference number: S65693
 A:Accession: S65693
 A:Molecule type: mRNA
 A:Residues: 1-392 <BAR>
 A:Cross-references: EMBL:U12569; NID:g607911; PIDN:AAB60354.1; PID:g607912
 R:Bare, L.A.; Mansson, E.; Yang, D.

FEBS Lett. 354, 213-216, 1994
 A:Title: Expression of two variants of the human mu oploid receptor mRNA in SK-N-SH cell
 A:Reference number: S51215; MUID:95046336
 A:Accession: S51216
 A:Molecule type: mRNA
 A:Residues: 387-392 <BAW>
 C:Superfamily: vertebrate rhodopsin

Query Match 16.3%; Score 357; DB 2; Length 392;
 Best Local Similarity 27.8%; Pred. No. 8.5e-23;
 Matches 100; Conservative 65; Mismatches 125; Indels 70; Gaps 12;

QY 11 SWIQOKLEDPQKHLNSTEYLAFLGPRRSHF-----FLPVSVVY 53
 DB 30 SWV-----NSHLDGDLSDPCGRNRDLGGRLDSCPTGSPSMATTAIALIS 78
 QY 54 PLFVGVGNVLVCLVLIHQAMKTPNYVLFSLAVSDLLVLLGMPLE-VYEMMRNYPE 112
 DB 79 IVCVGLFENFLVMYIVRYTKMKATNIIYIFNLALDALATST-LPFGSVNYLNGTWP- 136
 QY 113 LGPVCYKRTALFEVVCASISITTVSEVERVALHPRAKLOSTRRAALILGIWG 172
 DB 137 -FGTILCKIVISIDYNNMFTSIFTLCTMSVDRIYAVCHPKALDFRTPRNAKLIINCMNI 195
 QY 173 FSVLFLPMTSIRKIFKHFHPNGSLVPGSATCTVI--KPMIYNFIQVTSFLF-YLLPM 229
 DB 196 LSAIGLPLPMFATITTKYR-----QGSIDCTLTFESHPTWYENLKIICVFIFAFIMPY 247
 QY 230 TVISVLVYLMALRLKKDKSL-ADENANIQRCRKSVMKMLFVLVFAICMAPPHIDR 288
 DB 248 LITTCYGLMILRLKSVRMISGSKERDLNR-----ITRMVLVVAVFVTCWPIHIYV 302
 QY 289 LPFSFPE-----EMESLAAVFNVLHVVGVEFYLLSSANPIIYNLSRFOAFQ 339
 DB 303 IKAALVTIDETTFQTVSMHFCLIA-----LGYNSCLNPLVLAFLDENFRRCFR 350

RESULT 15

A57510
 mu oploid receptor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 24-Nov-1999
 C:Accession: A57510; I48665; S66513; I49300
 J. Krautman, D.L.; Keith Jr., D.E.; Anton, B.; Tian, J.; Magendzo, K.; Newman, D.; Tran, T.
 J. Biol. Chem. 270, 15877-15883, 1995
 A:Title: Characterization of the murine mu oploid receptor gene.
 A:Reference number: A57510; MUID:95318184
 A:Accession: A57510
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-398 <KAU>
 A:Cross-references: GB:U19380
 R.Mlin, B.H.; Augustin, L.B.; Felsheim, R.F.; Fuchs, J.A.; Loh, H.H.
 Proc. Natl. Acad. Sci. U.S.A. 91, 9081-9085, 1994
 A:Title: Genomic structure analysis of promoter sequence of a mouse mu oploid receptor
 A:Reference number: I48665; MUID:94377496
 A:Accession: I48665
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-398 <RES>
 A:Cross-references: EMBL:U10561; NID:9555696; PIDN:AAB60673.1; PID:95565069
 R.Rosel, G.C.; Pan, Y.X.; Brown, G.P.; Pasternak, G.W.
 FEBS Lett. 369, 192-196, 1995
 A:Title: Antisense mapping the MOR-1 oploid receptor: evidence for alternative splicing
 A:Reference number: I49300; MUID:95377399
 A:Accession: S66513
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-398 <ROS>
 A:Cross-references: EMBL:U26915; NID:91055230; PIDN:AAA61170.1; PID:91055231
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
 C:Genetics:

A:Gene: MOR-1
 A:Introns: 95/2; 213/1; 386/3
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmem

Query Match 16.3%; Score 357; DB 2; Length 398;
 Best Local Similarity 27.5%; Pred. No. 8.6e-23;
 Matches 105; Conservative 67; Mismatches 144; Indels 66; Gaps 14;

QY 37 CGPRR-----SHFLLP-----VSVVYPIPLFVGVGNVLVCLVLIHQAMKTP 79
 DB 43 CGPNRTGLGSHSLCPTGSPSMVTAITTAIALSIVCVGLFENFLVMYIVRYTKMKTA 102
 QY 80 TNYVLSLAVSDLLVLLGMPLE-VYEMMRNYPELGPGCYKRTALFEVVCASISIT 138
 DB 103 TNYIIFNLALDALATST-LPFGSVNYLNGTWP--FGNIIKCKIVISIDYNNMFTSIFTL 159
 QY 139 TVSEVERVALHPRAKLOSTRRAALRIIGIWGFSVLSPMTSIRKIFKHFHPNGSLV 198
 DB 160 TMSVDRIYAVCHPKALDFRTPRNAKIYVNCNIISSAIGLPLPMFATITTKYR----- 211
 QY 199 PGSATCTVI--KPMIYNFIQVTSFLF-YLLPMVIVSVLYLMALRLKKDKSL-ADGC 254
 DB 212 QGSIDCTLTFESHPTWYENLKIICVFIFAFIMPYLITTCYGLMILRLKSVRMISGSKER 271
 QY 255 NANIQRCKRSMKMLFVLVFAICMAPPHIDRLEFSFPE-----EMESLAAYV 305
 DB 272 DRNLRR-----ITRMVLVVAVFVTCWPIHIYVILKALITIDETTFQTVSMHFCLIA--- 323
 QY 306 NLHVVGVEFYLLSSANPIIYNLSRFOAFQNY---ISSFRKOWHS-----QHPOLP 358
 DB 324 -----LGYNSCLNPLVLAFLDENFRRCRECIPTSSITIEQNSAIRINQREHP 374
 QY 359 PAQRNIFLTCHEFVELTEDIGP 380
 DB 375 STANTVDRTNHOLENLEAETAP 396

Search completed: April 22, 2002, 14:17:43
 Job time: 296 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: Apr11 22, 2002, 15:47:04 ; Search time 116.57 Seconds
(without alignments)
2521.819 Million cell updates/sec

Title: US-09-609-146-3

Perfect score: 1298
Sequence: 1 agggagagcctcagcctctgg.....ttcagagcctcctctc

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/1na/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/1na/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/1na/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/1na/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/1na/PCTUS.COMB.seq: *
6: /cgn2_6/ptodata/1/1na/Backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	149.6	11.5	1088	4	US-09-077-675A-6 Sequence 6, Appl
2	144.6	11.1	1063	4	US-09-077-675A-1 Sequence 1, Appl
3	134.8	10.4	1092	4	US-09-077-675A-15 Sequence 15, Appl
4	120.4	9.3	1122	4	US-09-077-675A-9 Sequence 9, Appl
5	117.6	9.1	836	4	US-09-077-675A-11 Sequence 11, Appl
6	112.4	8.7	1029	4	US-09-077-675A-4 Sequence 4, Appl
7	112.4	8.7	1161	1	US-08-086-439C-2 Sequence 2, Appl
8	112.4	8.7	1161	1	US-08-434-877-2 Sequence 2, Appl
9	112.4	8.7	1367	3	US-08-475-742-3 Sequence 3, Appl
10	112.4	8.7	1370	1	US-08-056-051-1 Sequence 1, Appl
11	112.4	8.7	1370	1	US-07-928-611-17 Sequence 17, Appl
12	112.4	8.7	1370	2	US-08-487-811A-17 Sequence 17, Appl
13	112.4	8.7	1370	4	US-09-060-694-11 Sequence 11, Appl
14	112.4	8.7	1370	5	PCT-US93-07370-17 Sequence 17, Appl
15	112.4	8.7	1466	1	US-08-056-051-3 Sequence 3, Appl
16	112.4	8.7	1466	1	US-07-928-611-19 Sequence 19, Appl
17	112.4	8.7	1466	4	US-08-487-811A-19 Sequence 19, Appl
18	112.4	8.7	1466	4	US-09-060-694-19 Sequence 19, Appl
19	112.4	8.7	1466	5	PCT-US93-07370-19 Sequence 19, Appl
20	112.4	8.7	1610	1	US-08-056-051-5 Sequence 5, Appl
21	112.4	8.7	1610	1	US-07-928-611-21 Sequence 21, Appl
22	112.4	8.7	1610	2	US-08-487-811A-21 Sequence 21, Appl
23	112.4	8.7	1610	4	US-09-060-694-21 Sequence 21, Appl
24	112.4	8.7	1610	5	PCT-US93-07370-21 Sequence 21, Appl
25	110	8.5	1529	3	US-08-838-876A-3 Sequence 3, Appl
26	110	8.5	1529	4	US-09-472-880-3 Sequence 3, Appl
27	109.2	8.4	3129	4	US-09-077-675A-14 Sequence 14, Appl

28	97.6	7.5	1575	3	US-08-858-876A-1 Sequence 1, Appl
29	97.6	7.5	1575	4	US-09-472-880-1 Sequence 1, Appl
30	97.2	7.5	283	4	US-08-993-088A-4 Sequence 4, Appl
31	93.2	7.2	1164	4	US-08-993-088A-6 Sequence 6, Appl
32	88.8	6.8	1601	1	US-08-722-001-7 Sequence 7, Appl
33	88.8	6.8	1997	1	US-08-722-001-27 Sequence 27, Appl
34	88.8	6.8	2004	1	US-08-722-001-11 Sequence 11, Appl
35	88.6	6.8	1639	1	US-08-334-698-5 Sequence 5, Appl
36	88.6	6.8	1639	1	US-08-228-932-5 Sequence 5, Appl
37	88.6	6.8	1639	1	US-08-468-939-5 Sequence 5, Appl
38	88.6	6.8	1639	2	US-08-406-855A-5 Sequence 5, Appl
39	88.6	6.8	1639	2	US-08-722-190-5 Sequence 5, Appl
40	88.6	6.8	1639	3	US-08-724-354-5 Sequence 5, Appl
41	88.6	6.8	1639	3	US-09-206-899-5 Sequence 5, Appl
42	88.6	6.8	1639	5	PCT-US95-04203-5 Sequence 5, Appl
43	87.6	6.7	1342	3	US-08-832-399-1 Sequence 1, Appl
44	87.6	6.7	1342	4	US-09-372-498-1 Sequence 1, Appl
45	87	6.7	1205	1	US-08-417-103-13 Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-09-077-675A-6
Sequence 6, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pal, Lee-Yuh
APPLICANT: Feigner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Peng, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1088 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-077-675A-6
Query Match 11.5%; Score 149.6; DB 4; Length 1088;

Best Local Similarity 51.4%; Pred. No. 8,8e-28;
Matches 450; Conservative 0; Mismatches 414; Indels 12; Gaps 4

[illegible]

RESULT 2
US-09-077-675A-1

Sequence 1, Application US/09077
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pal, Lee-Yuh
APPLICANT: Feylghner, Scott C
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung

APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY

NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:

ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave
CITY: Rahway

CITY: KANAWAY
STATE: NJ
COUNTRY: USA

COMPUTER READABLE FORM
ZIP: 07065-0900
COUNTRY: USA

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FastSEO for Windows Vers

CURRENT APPLICATION DATA
APPLICATION NUMBER:
ATTENDING PHYSICIAN: 3

FILED DATE: 3-JUN
CLASSIFICATION:
PRIOR APPLICATION DATA

FROM AFFILIATE
APPLICATION
FILING DATE:

ATTORNEY/AGENT INFORMATION
NAME: Cocuzzo, Anna L.

REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P

TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273

TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR ID NO:

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1063 base pairs

LENGTH: 1000 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single

TOPOLOGY: linear
 MOLECULE TYPE: cDNA

9-077-675A-1

11.18; Score 144.6; DB 4; Length 1063

Method	Local Similarity	Conservative	Mismatches	Indels	Gaps
1	51.38	0	404	12	4

OY	167	cgatcctgtgtagtcatgtatgtccaaatttttgttgggtgggactatgtggcaatgcctcgtgtgtg	228
Db	97	CGTCAACCCGACACCTGCGCGGGCTTCTGTTGGTGGATATCCGGGGCAACCTGTCACAGAT	156
OY	227	ccctgtgtgtcttcgaagcaaccagcgtctatagaagaagcccaacactactactctctcaagct	286
Db	157	GCTGGTACTGTACAGCTTCCGCGAGATGCGACACACACCACTCAACTCTACCTGTGTACACAT	216
OY	287	ggcggctctctgaagccctcgtgtcccgctctcttgatgtccccctgtgaagtctatagaatgtg	346
Db	217	GGCCTTCTCGAAGCTAC---TCATCTTCTCTGTGATGATGCCCTTCACACTCTTCCGCCCTTG	273
OY	347	ggcgaactaacctcttcttgtctcgggcccgttgagctgtgctactcaagaacggcctcttga	406
Db	274	GCAGTACCGGGCTTGGAACTTGGCAACCTGCTTGCAAACTTTCCAGTTTCGTTAAGCA	333
OY	407	gaacgtgtgtctctgccttcacatccatcaagatcaacacacgctcaagcttgagcgtcatcgtgc	466
Db	334	GAGCTGCACCTACCGGCACAGTGTCTACCATATACCGCGCTGAGCTCGAGGCGCTTACTTCGC	393
OY	467	catctctaacaccgtctccgcgcgaacatgcagagagaccgcgagcgcggagccttcagatccct	526
Db	394	CATCTGCTTCCCGCTGGCGGGGCCAAGTATGTGTATACCAAGGCGCGGTTAAAGCTGTGTCAT	453
OY	527	cggagatcgtcttgggggcttctccgtgtctcttctcccttcccaacacagatcatcatgtgcat	586
Db	454	CTGTGTATCTGGGGCGTGGGCTTCTGTGAGGCGCGGGCCCATCTTTCGTGTGTGGAGAT	513
OY	587	caagttccactactctcccaatgggtcccttgttcccaagttcggcaacctgtacgtgcat	646

```

Db 514 GGAGCATGTAACGGCAGACGACCTCGGGACACCAAGAGTCCGCGCCACGAGATTCCG 573
Oy 647 caagcccatgtagtatacaatttcaatcacaagtcacccctccctcaatcaactcct 706
Db 574 CGTGGCCTCGGGGCTGCTTACCGTCAATGGTCTGGGGTCAAGTGTCTTCT--TCT 630
Oy 707 ccccatgactgcatcagtcctcctactacactcagtcacactcaagaataagaacaa 766
Db 631 GCGTGTCTTCTGCTCAGCTACGCTGTCTATAGCCTCATCGGACAGAACCTTGGGGAGAA 690
Oy 767 atcctctgagcagatgaaaggaatgcaaatltcaagaacctgaganaaatcagtcaa 826
Db 691 GCGCGCGGAGGGGGGCTGGCTCCCTCGCTCAGGGACAGAACCAACAAACCCGTG-- 748
Oy 827 caaagatgctgttctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtct 886
Db 749 -AAATGCTGGGTGAGTGGTGTGGCTTTCATACTCTGCTGCTGCTGCTGCTGCTGCT 807
Oy 887 ccga--cctctctcagcttgtagagagtagagtagaattccctgtgctgtgtctaa 943
Db 808 GCGATATTATTTCCTCAATCTCTGAGCCTGCTGTGAGAGATTGCTCAGATCAGCCA 867
Oy 944 cctgtccatgtagtgcagtgctctctctctcactcagctcagctgtaaccccatat 1003
Db 868 ATACTGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 927
Oy 1004 ctataacctactgtc 1018
Db 928 GTACAACATCATGTC 942

```

RESULT 3

US-09-077-675A-15

Sequence 15, Application US/09077675A

Patent No. 6242199

GENERAL INFORMATION:

APPLICANT: Pal, Lee-Yuh

APPLICANT: Reigner, Scott C.

APPLICANT: Howard, Andrew D.

APPLICANT: Pong, Sheng-Shung

APPLICANT: Van Der Ploeg, Leonardus H.T.

TITLE OF INVENTION: RECEPTOR ASSAY

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Merck & Co., Inc.

STREET: P.O. Box 2000, 126 E. Lincoln Ave.

CITY: Rahway

STATE: NJ

COUNTRY: USA

ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/077,675A

FILING DATE: 3-JUN-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Cocuzzo, Anna L.

REGISTRATION NUMBER: 42,452

REFERENCE/DOCKET NUMBER: 19590P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 732-594-1273

TELEFAX: 732-594-4720

TELEX:

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

```

; LENGTH: 1092 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-077-675A-15

```

Query Match 10.4%; Score 134.8; DB 4; Length 1092;
 Best Local Similarity 50.2%; Pred No. 3.7e-24;
 Matches 446; Conservative 0; Mismatches 427; Indels 15; Gaps 4;

```

Oy 167 cgtgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 226
Db 132 CGTACACGGGACACTGGGCTGCGGCTTTCGTTGGTGGGCACTGACGAACTGCTCAT 191
Oy 227 cctgtgtatctgtcagaccaggctatgaagagcccaactactactcctcact 286
Db 192 GCTGTGTGTGTCCGCTTCCGGGAGCTGCGCACACACCAACCTTACCTGTCCAGCAT 251
Oy 287 ggggtgtctgtgctcctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 346
Db 252 GGCCTTCTCGGATCTGCTCATCT--TCTGTGATGCGGCTGAGCTGTGCGCTGTG 308
Oy 347 ggcgaactaacctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 406
Db 309 GCGATACCGGCGCTGGAAGTCTGCGGACCTGCTGCAAACTTTCAGTTGTGAGCA 368
Oy 407 gacgtgtgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 466
Db 369 GAGCTGACACTACGCGACGCGCTCTACATACCGGCTGAGGCTGAGCGCTACTTCCG 428
Oy 467 catcccaaacccgtgtcgggccaactgtcagagcagcgcgcccgccctcaagatcct 526
Db 429 CATCTGCTTCTCTCTCGGAGCAAGGTGTGTGCTTAAAGGCGGTGAGCTGTGCTAT 488
Oy 527 cggcatcgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 586
Db 489 CTTTGTATCTGTGGGCTGTGCTTTCGAGCGGGGGCCCATTTGCTGTGGTGGCGT 548
Oy 587 caagttccactacttcccaatgtgtccctgtgtcccaagttcggcaactgtgaagtc 646
Db 549 GGAGCAGAAACGGCAGCAGATCCCGGAGACAAAGATGCGCGCACGAGTTCCG 608
Oy 647 caagcccatgtgtatctaaattcaatcaatcaatcaatcaatcaatcaatcaatc 706
Db 609 TGTGCGCTGTGGGCTGTACCGCTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTG 665
Oy 707 ccccatgactgtatcagtcctcactactcactcactcactcactcactcactcactc 766
Db 666 ACCGGTCTTCTGCTCAGTGTGCTTACAGTCTCAGGAGAGAACCTATGCGGAG-- 722
Oy 767 atcctctgagcagatgaaaggaatgcaaatltcaagaacctgaganaaatcagtcaa 826
Db 723 ---ACGCGGAGATGCAAGCGGTGGGCGCTCGCTCCGGGACCAACCAAGCAGCACT 779
Oy 827 caaagatgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 886
Db 780 GAAGATGCTGTGCTGTGGTGTGTTGCTTTCATCTGCTGTGCTGCTGCTGCTGCTG 839
Oy 887 ccga--cctctctcagcttgtagagagtagaattccctgtgctgtgtgtca 943
Db 840 AAGATACCTCTTTCGAAGTCTTGCAGCGCTGCTGTGGAATGTGCAAGTCAAGTCA 899
Oy 944 cctgtccatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1003
Db 900 GTACTGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 959
Oy 1004 ctataacctactgtctgcgcgtccagcagcagcagcagcagcagcagcagcagcag 1051
Db 960 GTACAACATCATGTCCAGAAGTACCGGGGTGCGAGTTCCTCAACTGCT 1007

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

	Query Match	Score	120.4	DB 4	Length	1122				
	Best Local Similarity	52.3%	Pred. No.	1.3e-20						
	Matches	290	Conservative	0	Mismatches	261	Indels	3	Gaps	1
Oy	167	cgtgcctgtgtgtatgtatgcacatttcttggtggggatcgtgcaatgtcctgtgtg	226							
Db	387	cgtacagcaccctgcgtgtgcactcttcgtgggtatcgcctggcaacctctcacact	446							
Oy	227	ccgtgtgattctgcagcacacagactatgaaagcccaactactactcttcagcct	286							
Db	447	gctgtgtgtgtgcgcgccttcgcggagctggcagcacacacacacactcttacctgtccagcatt	506							
Oy	287	ggcgtctcttgaccctccgtgtccctctcttgtaatgccccctggaaagtctatgaagtgtg	346							
Db	507	gcccttcttccgact---cttcctcatcttctcttcgcacaccccttggacactctgttcgacctctg	563							
Oy	347	ggcgaacatacccttctctgtgtcgggcctgtggcctgctcttccaagaacggcctcttga	406							
Db	564	gcagtaacgcgccttcgaaacttgcgggacactctcttcgcaaaccttcccaattcgtcagtga	623							
Oy	407	gaacgtgtgtcttgccctcatctctcaagaatacaaccctgcagaagcttgaaagcttactgtgc	466							
Db	624	gagctgcacacttgcgcacacgtctctcaccatcaccagcctgcagcgttcgacacgcttactctcc	683							

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	9.18;	Score 117.6;	DB 4;	Length 836;
Best Local Similarity	50.38;	Pred. No. 5.6e-20;		
Matches 372;	Conservative	0;	Mismatches 359;	Indels 9; Gaps 3;

[illegible]

```

RESULT 14
PCT-US93-07370-17
: Sequence 17, Application PC/TUS9307370
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses
: NUMBER OF SEQUENCES: 22
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US93/07370
: INFORMATION FOR SEQ ID NO: 17:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1370 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: 5'UTR
: LOCATION: 1..103
: FEATURE:
: NAME/KEY: 3'UTR
: LOCATION: 1268..1370
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 104..1267
PCT-US93-07370-17

```

Query Match	8.7%	Score 112.4	DB 5	Length 1370
Best Local Similarity	53.4%	Pred. No. 1.3e-18		
Matches 236	Conservative 0	Mismatches 206	Indels 0	Gaps 0
Oy	174	gtgtgtatgtgtccaaattttgtgtgtgtgtgtatgtgcaatgtctgtgtgtgtgtgt	233	
Db	218	gt	277	
Oy	234	attctgcagcaccaggtctatgtagaagcgcaccacaactatactcttcagccttggtgtc	293	
Db	278	gt	337	
Oy	294	tctgtacgtctgt	353	
Db	338	gt	397	
Oy	354	tacctcttctgttctgt	413	
Db	398	gt	457	
Oy	414	gt	473	
Db	458	gt	517	
Oy	474	caaccgttctgt	533	

Db	518	GTGGCGGTGGCGCTACAAACGGCAGGGTGTGGAGCGCGCGGACAGTGTGTGTCTATGTGGCGCC	577
Oy	534	gtctggaggctctctctgtgcctctctctcccttcgccaacacacagatcatgcatcaaatgc	593
Db	578	ACGTGGCTGCTGTCCGGCGCGGTGGCGCGCCCTACTGTGGGCTTCAAACGACGTGGCG	637
Oy	594	cactactctccccaatgggtgcc	615
Db	638	GGCGCGGACGCCGCGGTGTGGCG	659
Oy			

```

RESULT 15 ~-3
US-08-056-051-3
: Sequence 3, Application US/08056051
: Patent No. 5516683
: GENERAL INFORMATION:
: APPLICANT: Grandy, David K
: APPLICANT: Bunzow, James R
: APPLICANT: Civelis, Olivier
: APPLICANT: Van Tol, Hubert H.-M.
: TITLE OF INVENTION: A No. 5516683el Human Dopamine Receptor and uses
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Allegretti & Witcoff, Ltd.
: STREET: 10 South Wacker Drive, Suite 3000
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/056.051
: FILING DATE: 19930429
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: No. 5516683nan, Kevin E
: REGISTRATION NUMBER: 35,303
: REFERENCE/DOCKET NUMBER: 90,1092-C
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-715-1000
: TELEFAX: 312-715-1234
: TELEX: 910-221-5317
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1466 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: 5'UTR
: LOCATION: 1..103
: FEATURE:
: NAME/KEY: 3'UTR
: LOCATION: 1364..1466
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 104..1363
US-08-056-051-3

```

	Query Match	8.7%	Score 112.4	DB 1	Length 1466
	Best Local Similarity	53.4%	Prod. No. 1.3e16		
	Matches 236	Conservative	0	Mismatches 206	Indels 0
					Gaps 0
Qy	174	gtggtatggtcccaatttctgtgtgtggggcattcgtgcaatgccctgtgtgctcgtgtg	233		
db	218	gtggggggcgtcctctcattcggccgcggcgtcggggcgaactccctctctggccgggac	277		

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 14:13:46 ; Search time 2258.39 Seconds
(without alignments)
6176.090 Million cell updates/sec

Title: US-09-609-146-3

Perfect score: 1298
Sequence: 1 agggagagctcagcctctg.....ttcagagctgactctctc 1298

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 537289281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estcpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_est1:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_dln:*
18: em_gss_pro:*
19: em_gss_tod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	378	29.1	555	10	BE751626 203809 MA
2	339	26.1	929	11	BE871983 601448032
3	302	23.3	414	11	M45474 yv59b04.r1
4	299.2	23.1	465	10	BE751310 203132 MA
5	211.2	16.3	941	13	CNS036WH
6	123.4	9.5	450	11	RL3890
7	116.4	9.0	959	13	CNS04GXP
8	113	8.7	379	11	HL1359
9	107.8	8.3	817	13	CNS01XJ2
10	107	8.2	500	11	BE603623
11	95.4	7.3	448	11	BF323227
12	95.2	7.3	975	13	CNS039RV

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
1	BE751626	203809 MARC 280V Bos taurus CDNA 5', mRNA sequence.	BE751626	1	GI:10165618	EST.	Bos taurus	Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.M., Rohrer, G.A., Chitko-McKown, C.G., Pette, G., Holt, I., Karanycheva, S., Liang, F., Quackenbush, J. and Keefe, J.W.	Sequence evaluation of four pooled-tissue normalized bovine CDNA libraries and construction of a gene index for cattle	Genome Res. 11 (4), 626-630 (2001)	21180013	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing. Bases called and alt-trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options. PCR Primers FORWARD: AGGAACACGCTATGACCAT BACKWARD: GTTTCACGTCACGACG Plate: 44 row: 0 column: 6	

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
1	BE751626	203809 MARC 280V Bos taurus CDNA 5', mRNA sequence.	BE751626	1	GI:10165618	EST.	Bos taurus	Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.M., Rohrer, G.A., Chitko-McKown, C.G., Pette, G., Holt, I., Karanycheva, S., Liang, F., Quackenbush, J. and Keefe, J.W.	Sequence evaluation of four pooled-tissue normalized bovine CDNA libraries and construction of a gene index for cattle	Genome Res. 11 (4), 626-630 (2001)	21180013	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing. Bases called and alt-trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options. PCR Primers FORWARD: AGGAACACGCTATGACCAT BACKWARD: GTTTCACGTCACGACG Plate: 44 row: 0 column: 6	

FEATURES Seq primer: ATTAGTGACACTATAG.
Location/Qualifiers
1. 555
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 2B0V"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPOR6; Site.1: XbaI; Site.2: XhoI;
Library made from pooled tissue from testis, thymus,
semilendonsus muscle, longissimus muscle, pancreas,
adrenal, and endometrium."

BASE COUNT 100 a 165 c 134 g 156 t
ORIGIN

Query Match 29.1%; Score 378; DB 10; Length 555;
Best Local Similarity 86.5%; Pred. No. 1.6e-81;
Matches 417; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 1 agggagagcctcagccttgatttaatgctcagggatgataaaactcagaatgctcct 60
DB 74 AGAGAGACCTCAGAGCTTGTGATTTTATACAGATGATAAATCATGATGTTCTTCT 133

QY 61 ggaactacagcagaactagaagatccattccagaacacactgaagacagcaggagt 120
DB 134 GGATGTACCAACAGAGATGATAAATGATCTTCAAGAAATACCTGAGACACAGATGACT 193

QY 121 atctgagcctcctcctcctgagacccctcgagcagcactcctcctcccgctcgtgtgtc 180
DB 194 ACCTAGCTTGTGCTCGCGGCTGCTGCGAGCACCCTTCTCCCGTGACTCGGCTGT 253

QY 181 atgtgcaattttgt 240
DB 254 ATGCTCTGATTTTGT 313

QY 241 agcagcagcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 300
DB 314 GGCACAGCAGATGAGAACCCACCAATTAATCTCTTCTGAGCTTGTGTGTGTGTGTGT 373

QY 301 tctctgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 360
DB 374 TCTCTAGCTTGT 433

QY 361 tcttctgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 420
DB 434 TCTCTGT 493

QY 421 cctccatcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480
DB 494 CCTCATCTCAGCGTCACGTCACGTCAGCGTCGATGTATGTATGTATGTATGTATGTAT 553

QY 481 tc 482
DB 554 TC 555

RESULT 2
BE871983 929 bp mRNA EST 20-OCT-2000
LOCUS BE871983
DEFINITION 601448032F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3852151 5',
mRNA sequence.
ACCESSION BE871983
VERSION BE871983.1 GI:10320759
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 929)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ARCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.lnl.gov
plate: L1AM9573 row: p column: 08
High quality sequence start: 11
High quality sequence stop: 602.
Location/Qualifiers
1. 929
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:3852151"
/clone_lib="NIH-MGC-65"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: PCMV-SPOR6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies."

BASE COUNT 186 a 252 c 228 g 263 t
ORIGIN

Query Match 26.1%; Score 339; DB 11; Length 929;
Best Local Similarity 86.1%; Pred. No. 5.7e-72;
Matches 516; Conservative 0; Mismatches 65; Indels 18; Gaps 12;

QY 1 agggagagcctcagcct-tgattttaatgctcagggatgataaaactcagaatgctcct 59
DB 107 AGGGAGAGCTCAGAGCTTGTGATTTTATACAGATGATAAATCATGATGTTCTTCC 166

QY 60 tgaactacagcagaactagaagatccattccagaacacactgaagacagcaggagt 119
DB 167 TGGATGTACCAAG -AGAAATCTACAGATTCATTCAAAACACCTGAGACAGCAGAGAG 225

QY 120 tatctgagcct-ttctctcctgagacccctcgagcagcactcctcctcccgctgtgtgtc 178
DB 226 TATCTGAGCTTGT 285

QY 179 gtatgtgcaattttgt 238
DB 286 GTATGTGCAATTTATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 345

QY 239 gcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 297
DB 346 GCAGCAGCAGGCTATGAGAACGCCACCAA -TACTACTCTTCTGAGCTTGTGTGTGTGT 404

QY 298 accctcctg--tctctgctccttgaatgagccttgaggtc--atgagatgtgtgtgtgt 352
DB 405 ACCTCTGT 464

QY 353 ctacccttctgt 409
DB 465 CTACCCCTTTTGT 524

QY 410 cgtgtgtcgtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 469
DB 525 CGT 582

QY 470 cctacacccgttccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 529
DB 583 CTTACACCCGTT -GGCCCAAAATGAGAGACACCCGGGTGCCGCTTACAGGATTCGCG 641

QY 530 catcgtcgtgagcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 588
DB 642 AATTCCTGT 698

```

RESULT 3
LOCUS N45474 414 bp mRNA EST 13-FEB-1996
DEFINITION y59b04.r1 Soares,multiple_sclerosis_2nbhmsp Homo sapiens cDNA
clone IMAGE:277807 5', mRNA sequence.
ACCESSION N45474
VERSION N45474.1 GI:1186640
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 414)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,
M., Hultman,M., Kucuba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevisan,E., Waterston,
R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: T7
High quality sequence stop: 264.
Location/Qualifiers
1. 414
/organism="Homo sapiens"
/db_xref="GDB:3896183"
/db_xref="taxon:9606"
/clone="IMAGE:277807"
/clone_1lb="Soares_multiple_sclerosis_2nbhmsp"
/sex="male"
/tissue_type="multiple_sclerosis_lesions"
/dev_stage="Age 46"
/lab_host="DH10B (ampicillin resistant)"
/notes="Vector: pT73D (Pharmacia) with a modified
polylinker V-type; phagemid; Site_1: Not I; Site_2: Eco RI
; 1st strand cDNA was primed with a Not I - oligo(dT)
primer 15'
TGTACCATCTGAAGTGGAGCGGCGCCATTTTTTTTTTTTTTTT 3'}.
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M. Fatima Bonaldo. RNA from 4 multiple sclerosis
lesions from one patient was kindly provided by Dr. Kevin
G. Becker (NINDS/NIH)."
BASE COUNT 113 a 123 c 71 g 105 t 2 others
ORIGIN
Query Match 23.3%; Score 302; DB 11; Length 414;
Best Local Similarity 100.0%; Pred. No. 4.7e-63;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 997 ccaataataaactctgtctgcgcgctccagcagcattccagaatgtatctctt 1056
|||||
Db 1 CCATTATCTATACTACTCTCTCGCCGCTTCCAGCAGCATTCAGATATGATCTCTT 60
QY 1057 ctttcacaacagctggcactccacagcatgacacagttgcacactgcccagcgagaca 1116
|||||
Db 61 CTTTCACAAACAGTGACACTCCAGCATGACCCACAGTGGACACTGCCAGCGGAGACA 120
QY 1117 tctctctgacagaatccactttgtgagctgcagcgaagatatagttcccaattcccat 1176
|||||
Db 121 TCTTCTGACAGAAATGCCACTTTGTGAGAGCTGACGCAAGATATAGTCCCAATTCCTCAT 180

```

```

QY 1177 gtcaatcatcagcacaactctcaactctcccaacagccctctctagtgaacagatgcaa 1236
|||||
Db 181 GTCAATCATCAGCACAACACTCTCACTCCCAACAGCCCTCTCTAGTGAACAGATGCAA 240
QY 1237 gaacaacatcaaaagcttccactttaaacaacactgaattcttccagctgactctc 1296
|||||
Db 241 GAACAACATCAAAAGCTTCACTTTAACAACAACTGAATTTCTTTCAGACTGACTCTCC 300
QY 1297 tc 1298
||
Db 301 TC 302
RESULT 4
LOCUS BE751310 465 bp mRNA EST 25-APR-2001
DEFINITION 203132 MARC 2B0V Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE751310
VERSION BE751310.1 GI:10165239
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 465)
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrnenkrug,S.C., Bennett,
G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G.,
Perte,G., Holt,I., Karanycheva,S., Liang,F., Quackenbush,J. and
Keeler,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
genome Res. 11 (4), 626-630 (2001)
21180013
CONTACT: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross_match with the -mnscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACACCTATGACCAT
BACKWARD: GTTTCACAGTCACGACG
Plate: 43 row: O column: 6
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers
1. 465
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_1lb="MARC 2B0V"
/tissue_type="pooled"
/lab_host="DH10B"
/notes="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from testis, thymus,
semitepidonous muscle, longissimus muscle, pancreas,
adrenal, and endometrium."
BASE COUNT 88 a 129 c 113 g 135 t
ORIGIN
Query Match 23.1%; Score 299.2; DB 10; Length 465;
Best Local Similarity 85.2%; Pred. No. 2.3e-62;
Matches 334; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 1 agggagagctcagcgcttgatatttaatgcaaggatggaanaactcagaatgctctc 60
|||||
Db 74 AGAGGACGCTCAGGCTCTTTGATTTATATACAGAGGAGAAACATGAAATGTTTCTT 133
QY 61 ggaatctacagcagaactagaagatcattccagaacacccctgaacagacgaggaagt 120

```

Query Match	16.3%	Score 211.2;	DB 13;	Length 941;
BASE COUNT	197 a	277 c	327 g	132 t
ORIGIN	8 others			
FEATURES	<p>source</p> <p>1. 941</p> <p>/organism="Tetraodon nigroviridis"</p> <p>/db_xref="taxon:99883"</p> <p>/clone="216F22"</p> <p>/clone_1lb="G"</p> <p>/note="Genoscope sequence ID : COAG216DC11npl-end : T7"</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p> <p>This sequence is a single read and was generated as part of a large</p> <p>scale clone and sequencing project of the Tetraodon nigroviridis</p> <p>genome. For more information, please take a look at</p> <p>http://www.genoscope.cns.fr/Tetraodon.</p>			
REFERENCE	<p>Direct Submission</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
AUTHORS	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
JOURNAL	<p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p>			
REFERENCE	<p>Submitted (12-APR-2000) to the EMB</p>			

[illegible]

Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.M., Rohrer, G.A., Chitko-McKown, C.G., Pettea, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and Keeler, J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle

Genome Res. 11 (4), 626-630 (2001)

21180013

JOURNAL MEDLINE

COMMENT

Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -mismatch 18 and -mismatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCAGTCAGCAGC

Plate: 49 row: I column: 6

Seq primer: ATTTAGTGACCTATGAC.

FEATURES

Location/Qualifiers

1..500

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone_lib="MARC 3BOV"

/tissue_type="pooled"

/lab_host="DH10B"

/note="Vector: PCMV SPOR6; Site_1: XbaI; Site_2: XhoI;

Library made from pooled tissue from marrow, alveolar

macrophage, ovary, fetal semitendinosus muscle, and fetal

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

longissimus muscle."

DEFINITION

maab38a07.y1 NCI CGAP Brn63 Mus musculus cDNA clone IMAGE:3813037 5' similar to SW:NTF2_RAT Q63384 NEUROFENSIN RECEPTOR TYPE 2 ;, mRNA sequence.

ACCESSION

BF323227

VERSION

BF323227.1 GI:11272778

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 448)

AUTHORS

NCI/NINDS-CGAP

TITLE

http://www.ncbi.nlm.nih.gov/ncicgap.

JOURNAL

National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project

COMMENT

(CGAP/BRGAP), Tumor Gene Index

Unpublished (1998)

Other ESTs: maab38a07.x1

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Sequencing by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

MG1:1455149

Seq primer: -40RP from Gibco

High quality sequence stop: 416.

Location/Qualifiers

1..448

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone_lib="IMAGE:3813037"

/clone_lib="NCI-CGAP_Brn63"

/sex="female"

/dev_stage="10 weeks"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: brain; Vector: PCMV-SPOR6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.8 kb. Library constructed by Life

Technologies."

BASE COUNT

53 a 157 c 148 g 89 t 1 others

ORIGIN

Query Match

Best Local Similarity 7.3%; Score 95.4; DB 11; Length 448;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Dd	308	GCCCGACGTGACTCAGTGTCTGGGTCGCCTCTCTGAGGCCCTTGCCCTCCCATGGCGGGTTATTC	367
Oy	579	catgacatcaagttcca	595
Dd	368	ATGGGACAGAAACACGCA	384
RESULT_12			
CNS039RV			
LOCUS	CNS039RV	975 bp	DNA
DEFINITION	Tetradodon nigroviridis genome survey sequence T7 end of clone 008M03 of library G from Tetradodon nigroviridis, genomic survey sequence.		
ACCESSION	AL334282		
VERSION	AL334282.1	GI:7893427	
KEYWORDS	GSS; genome survey sequence.		
SOURCE	Tetradodon nigroviridis.		
ORGANISM	Tetradodon nigroviridis Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Aclinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontoidei; Tetradon.		
REFERENCE	1 (bases 1 to 975) Roest-Crollius,H., Jalllon,O., Dasilva,C., Fitzames,C., Fisher,C., Bouneau,L., Billault,A., Quettier,F., Saurin,W., Bernot,A. and Weissenbach,J. Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradodon nigroviridis		
AUTHORS	Unpublished		
JOURNAL	2 (bases 1 to 975)		
REFERENCE	Roest-Crollius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fitzames,C., Wlncker,P., Brotlier,P., Quettier,F., Saurin,W. and Weissenbach,J. Human gene number estimate provided by genome wide analysis using Tetradodon nigroviridis DNA sequence		
AUTHORS	Unpublished		
TITLE	3 (bases 1 to 975)		
JOURNAL	Genoscope.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (12-Apr-2000) to the EMBL/Genbank/DBJ databases		
TITLE	This sequence is a single read and was generated as part of a large scale clone-and-sequencing project of the Tetradodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetradodon.		
JOURNAL	Localion/Qualifiers		
COMMENT	1..975		
FEATURES	source		
	/organism="Tetradodon nigroviridis"		
	/db_xref="taxon:99883"		
	/clone="008M03"		
	/clone_11b="G"		
	/note="Genosope sequence ID : COBG008AG02LP1-end : T7"		
	BASE COUNT 164 a 332 c 295 g 179 t 5 others		
	ORIGIN		
Query Match	7.3%	Score 95.2;	DB 13; Length 975;
Best Local Similarity	54.0%;	Pred. No.1.2e-12;	
Matches 218:	Conservative 0;	Mismatches 183;	Indels 3; Gaps 1;
Oy	175	tgggtatgtgcgaatttcttggtggtggggtcattggcgaagtctctggtgtgcctggtga	234
Dd	231	TGGTGGCCGGCTCCCTCACCGCTGGTCACCGTATGGGCAACATCTGTGCATCGTGTCCA	290
Oy	235	tctctcagaagcaggactgatgaagcgcccaaccaactactacccttcgaagctggsgtgtc	294
Dd	291	TCAAGTCAACCGCAACTCTGCAGACGGTCAACAACACTACTTCTGTTCACCTCTGGCGTGG	350
Oy	295	ctgacctctctgttctctgtcctcttgaaatgccctctggaaggtctcataagatgttgcacaact	354
Dd	351	CGGACTGTATCATTCGGCCCTGTGCTCATCAACTGTACACCGCTMACATCGTAGTTGGCT	410
Oy	355	acccttctctgttcggcccgtggagctgactactcaagaagcgccctctttgagaccgtgt	414

QY	411	AC---	TGGCGCGTGGGGCCCGGNGTGTGGAGACCTTGGCTGGCGCTGGACACTGACGTGGA	467
QY	415	gcttcgctccatcctcctcagcatcacccgctcagcgtagcgctagctgagcctcctac	474	
Db	468	GCAGAGCCCTCCGTCGTGAGAACCTTCATCATCATCAGCTTGAGCCGCACTTCGCGTACCA	527	
QY	475	accgcttcgagccaaactgcagagcaccgcgagccggcctcagatccctcgagcagc	534	
Db	528	AGCCGCTCAGCTACACCGGGGAAGCGCCACACCAAGATGCGGGGCAATGATGATCGCGCTG	587	
QY	535	tcctgggcttcctcgctcctcctccctgcgcacacacagcaltc	578	
Db	568	CCTGGGTGCTGTCTTCATCATCTTCCTGGGGCGCGGCATCTCTTC	631	
RESULT 13				
CNS020M1/C				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				

Db 775 CATGACGATGCGGAGGACCTACGAGGCGCCWTCGTCGTCCTTACCTTACGAGCGTCGCGGAC 716
 Qy 372 cccggtggtgtactactcaagaagccctcttgagacggtgtgtctgctccatccctc 431
 Db 715 AACCTCTCTCTGTATACACACTTCTGTTTCGAGGCGGCGGACACGCCACATGTGG 656
 Qy 432 agcatcacacacgctgagcgctgagcgtcgtgagccatccctacacccgttcgagccaaa 491
 Db 655 DACATAGCCACACCTTDCCTTCAGAGCGCTACGTGGCCCTTTCGCCACCCCTCCGCTACAG 596
 Qy 492 ctgcagagacacccggtgagccgagccctcagatcctgcagctcgtcgtgagctctccgtg 551
 Db 595 GCCCTGGG---CGGAGGCGGACACMACGCCCTCMTGCGCTTGGCCCTKGTGGTGTGCGGG 539
 Qy 552 cctctctccctgcccacacacgcatcgtgcatcgaagttccactactctccc 605
 Db 538 CTGCTGGCCCTGCGCTGCTCATCGCACGAGGACGACGAGNCGACATCCCTC 485

RESULT 14

Locus CENS04P41 1079 bp DNA GSS 24-MAY-2000
 DEFINITION Tetraodon nigroviridis genome survey sequence PUC-ORI end of clone 125F12 of library G from Tetraodon nigroviridis, genomic survey sequence.
 ACCESSION AL300843.1 GI:8177862
 VERSION AL300843.1
 KEYWORDS GSS; genome survey sequence.
 SOURCE Tetraodon nigroviridis.
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphi; Acanthopterygii; Percomorphi; Tetraodontiformes; Tetraodontidae; Tetraodon.
 1 (bases 1 to 1079)
 Roest-Crollius,H., Jallion,O., Daaliva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
 Unpublished
 2 (bases 1 to 1079)
 Roest-Crollius,H., Jallion,O., Daaliva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Winkler,P., Brotlier,P., Queller,F., Saurin,W. and Weissenbach,J.
 Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
 Unpublished
 3 (bases 1 to 1079)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases
 This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.

FEATURES

Source 1..1079
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone="125F12"
 /clone.lib="G"
 /note="Genoscope sequence ID : C0BG125DC06SP1-end : PUC-ORI"

BASE COUNT 196 a 347 c 270 g 259 t 7 others
 ORIGIN

Query Match 73% Score 94.6; DB 13; Length 1079;
 Best Local Similarity 64.7%; Pred. No. 1.7e-12;
 Matches 139; Conservative 1; Mismatches 75; Indels 0; Gaps 0;

Qy 839 tgccttgcttagtgccttgcctatcgttgcgcccgttccacattgacgcgactctct 898
 Db 457 tcttctggtggtggtggttcttgcgctctgctgagccccccttttacgtagacgcttatgmg 516
 Qy 899 cagctctgtgagagagtgagtgaaatccctgcgtcgtgtgttcaacctgcgtcatgtgt 958
 Db 517 GAGCTGATAGGACGTCCTCCTCCGACGCGCAGCTGAGGGGCTCTGAGAGGTCTACGTCAT 576
 Qy 959 gtcagtgctctctctctacgtagctcgtgctcaacccattatctataactactgtc 1018
 Db 577 CTGTGAGATMTCTCTTACTGAGCTCGCGCGTCAACCCCAATCCCTTACAACTGATGTC 636
 Qy 1019 tccgcgtccagcagcatccagatgtgactc 1053
 Db 637 CACCCGCTTCAGGAAACGTTACGCCACGCTACCT 671

RESULT 15

Locus AK004891 2611 bp mRNA HTC 05-JUL-2001
 DEFINITION Mus musculus adult male liver cDNA, RIKEN full-length enriched library, clone:1300003P04, full insert sequence.
 ACCESSION AK004891
 VERSION AK004891.1 GI:12836417
 KEYWORDS CAP trapper.
 SOURCE Mus musculus (strain:C57BL/6J) adult male liver cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library
 clone:1300003P04.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE AUTHORS Carninci,P. and Hayashizaki,Y.
 TITLE High-efficiency full-length cDNA cloning methods in enzymology. 303, 19-44 (1999)
 JOURNAL 99279253
 MEDLINE 10349636
 PUBMED
 REFERENCE AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL 20499374
 MEDLINE 11042159
 PUBMED
 REFERENCE AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichipillary sequencer
 JOURNAL 20530913
 MEDLINE 11076861
 PUBMED
 REFERENCE AUTHORS

4 (bases 1 to 2611)
 THE RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 REFERENCE 5 (bases 1 to 2611)
 Adachi,T., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arikawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Horii,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J., Kojima,Y., Kono,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shingawa,A., Shikaki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,

TITLE
JOURNAL

Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUN-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGCGCCCACTCGATTGTTTGTGTTTATN 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGAGAGATCCAGATCAATTAATTAATAAACCCGCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOLR.

FEATURES

source

Location/Qualifiers
1. 2611

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/db_xref="MGI:104774"

/db_xref="MGI:1907258"

/clone="1300003P04"

/sex="male"

/tissue_type="liver"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

513..2060

/note="putative"

/codon_start=1

/protein_id="BAB23647.1"

/db_xref="GI:12836418"

/translation="MNPDLDTGHNTPAPAHGELKDANFTGNQSSNSTLPOLDVTR
AIVSGIVGAFILFAIVGNILVLSVACNRHLRTPNYIVNLAIDLSTVLPFS
ATLEVGVVLCRIFCDIWAADVLCSTASISIDRIGVRSLOYPLVTR
KAILALSVWLVSTVISIGPLGKMEKPEANDKEGVTEPFALESFYSFIPLAV
ILWVCRAVIVAKRTTKNLKAGVWKMENSKRLTRHSKNHEDTISTKAGNPR
SSIAVLEKESREKRAKTGLVYGCFTICMPTFIALPLGSLFSLKPPDAVKYF
WGYRNSCLNPITITPCSSKERRAFKRIIGCCRGRRRRRRRLGGCAVTRPMTG
GSLERSQSRKSDLDSDSGSCMSQRTLPSPSPGVLGRTQPPVELCAFPMKPGAL
LSLPEPPGRGRRLDSGLPTFTFLKLGEPSPTEGDASNGCDTTDLANGDPGFSNM
PLPGHF"

polyA_signal

2590..2595

polyA_site

/note="putative"

2611

/note="putative"

/note="putative"

BASE COUNT 535 a 776 c 740 g 560 t

ORIGIN

Query Match

Best Local Similarity 7.0%; Score 90.6; DB 12; Length 2611;

Matches 289; Conservative 0; Mismatches 309; Indels 3; Gaps 1;

QY 161 cctcccgtgtcgtgtgtgtatgtgccaatttctgtgtgtgtgtgtatgtgcaatgtcct 220

DB 647 CATCTCTGTGGCGCTGTGGCGCTTCATCTTGTGCGCATCGTGGCAACATCTT 706

QY 221 ggtgtcctgtgtatctgtcagcagcaggtatgaagaagccaccactactactctt 280

DB 707 ggtcatctctgt 766

QY 281 cagcctgt 340

DB 767 CAACCTGGCCATTGCTGACCTGCTGTTGAGCTTCACAGTCTGCCCCCTTCCGCTACCT 826

QY 341 gatgt 400

DB 827 GGAAGTGTCTGGCTAC--TGGGTGCTGGGGCGCATCTTCTGTGACATCTGGGACGGGT 883

QY 401 ctltgt 460

DB 884 TGATGTCTGT 943

QY 461 cgt 520

DB 944 CATTTGGGTGCGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1003

QY 521 gatcctgt 580

DB 1004 GGGCTCTCAGT 1063

QY 581 tggatcaaatgt 640

DB 1064 GAAGAAGCTGCGCCCAATGACGACGAAGAAGATGTGGGGTACAGAGAAGAACCTTACGC 1123

QY 641 ggtcatcaagccatgt 700

DB 1124 CCTCTTCT 1183

QY 701 cctcctcccatgt 760

DB 1184 CCGAGTCTACATGT 1243

QY 761 a 761

DB 1244 A 1244

Search completed: April 22, 2002, 17:16:21
Job time: 10955 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 15:17:21 ; Search time 3659.47 Seconds
(without alignments)
5851.486 Million cell updates/sec

Title: US-09-609-146-3

Perfect score: 1298
Sequence: 1 agggagagcagcagccttg.....ttcagagcctcctctc 1298

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pal:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rnd:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rnd:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1298	100.0	1298	9	AF272363	AF272363 Homo sapi
2	1298	100.0	1658	6	AX109234	AX109234 Sequence
3	1298	100.0	1658	6	AX109236	AX109236 Sequence
4	1291.6	99.5	1658	6	AX109238	AX109238 Sequence
5	1291.6	99.5	1658	6	AX109240	AX109240 Sequence
6	1248	96.1	1248	9	AB041228	AB041228 Homo sapi
7	1246.4	95.0	1249	9	AF292402	AF292402 Homo sapi
8	1239	95.5	1239	9	AF242874	AF242874 Homo sapi
9	1160	89.4	1594	6	AX109242	AX109242 Sequence
10	1160	89.4	1594	6	AX109244	AX109244 Sequence
11	802.8	61.8	1314	10	AB041229	AB041229 Rattus no
12	788.8	60.8	1188	10	AF242875	AF242875 Rattus no
13	752.8	58.0	215088	2	AC008571	AC008571 Homo sapi
14	723.8	55.8	801	6	AX147772	AX147772 Sequence
15	717.8	55.3	729	6	AX139107	AX139107 Sequence
16	346	26.7	1318	9	AF272362	AF272362 Homo sapi
17	295.2	22.7	1209	10	AF242873	AF242873 Rattus no
18	293.6	22.6	1239	10	AB038649	AB038649 Rattus no
19	281	21.6	828	9	HS06PCR1	HS06PCR1
20	281	21.6	168880	9	AC017104	AC017104 Homo sapi
21	270.6	20.8	1218	10	AF044602	AF044602 Mus muscu
22	159.8	12.3	75950	2	AC073449	AC073449 Homo sapi
23	149.6	11.5	1088	6	AR156353	AR156353 Sequence
24	149.6	11.5	1101	6	AX154584	AX154584 Sequence
25	149.6	11.5	1101	6	HS060179	HS060179
26	144.6	11.1	1063	6	AR156351	AR156351 Sequence
27	144.6	11.1	1101	4	SS060178	SS060178
28	135.4	10.4	1676	5	AF082210	AF082210 Spheeroid
29	135	10.4	1098	6	AX154586	AX154586 Sequence
30	134.8	10.4	1092	6	AR156357	AR156357 Sequence
31	134.8	10.4	1161	6	AX154581	AX154581 Sequence
32	134.8	10.4	1239	6	AX154589	AX154589 Sequence
33	134.8	10.4	1350	10	AB001982	AB001982 Rattus no
34	134.8	10.4	2040	9	AF034632	AF034632 Homo sapi
35	134.8	10.4	163284	2	AL137000	AL137000 Human DNA
36	134.8	10.4	341560	2	AL596304	AL596304 Homo sapi
37	126.8	9.8	1254	22	EL1480	EL1480 cDNA encodi
38	126.8	9.8	4131	9	HSNEURA	HSNEURA
39	120.4	9.3	870	9	HS060181	HS060181 Human growt
40	120.4	9.3	1122	6	AR156354	AR156354 Sequence
41	120.4	9.3	6787	9	AF369786	AF369786 Homo sapi
42	120.4	9.3	145593	9	AC069523	AC069523 Homo sapi
43	119.8	9.2	121652	34	AC008191	AC008191 Drosophill
44	119.8	9.2	140838	2	AC018176	AC018176 Drosophill
45	119.8	9.2	195212	3	AC007441	AC007441 Drosophill

ALIGNMENTS

RESULT 1
LOCUS AF272363 1298 bp mRNA PRI 23-OCT-2000
DEFINITION Homo sapiens neuromedin U receptor 2 (NMUR2) mRNA, complete cds.
ACCESSION AF272363
VERSION AF272363.1 GI:10946202
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Radatz,R., Wilson,A.E., Artymyshyn,R., Bonini,J.A., Borowsky,B.,
Botelju,L.W., Zhou,S., Kouranova,E.V., Nagorny,R., Guevarra,M.S.,
Dai,M., Lerman,G.S., Vayse,P.J., Branchet,T.A., Gerald,C.,
Forray,C. and Adham,N.

TITLE Identification and Characterization of Two Neuromedin U Receptors
Differentially Expressed in Peripheral Tissues and the Central
Nervous System
JOURNAL J. Biol. Chem. 275 (42), 32452-32459 (2000)

PUBMED 10899166
 REFERENCE 2 (bases 1 to 1298)
 AUTHORS Bonini, J.A., Reddatz, R., Wilson, A. and Borowsky, B.
 TITLE Direct Submission
 JOURNAL Submitted (25-MAY-2000) Target Discovery and Assessment, Synaptic
 Pharmaceutical Corporation, 215 College Road, Paramus, NY 07652,
 USA

FEATURES
 source Location/Qualifiers
 1..1298
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /map="5q31.1-q31.3"
 1..1298
 /gene="NMUR2"
 27..1274
 /gene="NMUR2"
 /note="G protein-coupled receptor"
 /codon_start=1
 /product="neuregulin 1 receptor 2"
 /protein_id="GI:10946203"
 /db_xref="GI:10946203"

gene
 CDS
 27..1274
 /gene="NMUR2"
 /note="G protein-coupled receptor"
 /codon_start=1
 /product="neuregulin 1 receptor 2"
 /protein_id="GI:10946203"

BASE COUNT 278 a 405 c 282 g 333 t
 ORIGIN

Query Match 100.0%; Score 1298; DB 9; Length 1298;
 Best Local Similarity 100.0%; Pred. No. 3.6e-259;
 Matches 1298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agggagagagcagagccttgatatttaattgcagagatgagaaacatcagaatgctctc 60
 Db 1 AGGGAGAGCTCAGGCGCTTGATATTTAATGTCAGGAGATGAAAAACTTCAGAAATGCTTCT 60
 QY 61 ggaactacagcagaagaactagaagatccagaagaacactgagacagcagagagt 120
 Db 61 GGATCTACAGCAGAAACTAGAAGATCTCAGAAACACCTGACAGCAGCAGAGT 120
 QY 121 atttgacctctctctgagagcctcgagcagacactctctcccccgtgtctgtgtg 180
 Db 121 ATCTGGCTCTCTCTGCGGACCTGCGGACCCACTTCTCTCCCGGTGTGTGTGT 180
 QY 181 atgtgcaaatcttctgt 240
 Db 181 ATGTGCCAATTTTGT 240
 QY 241 agcacaagcctatgaagaagcagcacaactactactctctcagcctgagcgtctcagc 300
 Db 241 AGCACAAGCCTATGAAGAAGCAGCAGCACTACTCTCAGCCTGCGGTCTGTGAC 300
 QY 301 tctctgtctcgtctctctgtgaatgccccctgtgaagtctatgtgaatgtgagcagaacccctc 360
 Db 301 TCTCTGTCTCTCTCTGTGAATGCCCTGTGAAGTGTGTGTGTGTGTGTGTGTGT 360
 QY 361 tctctgtcagcagcctgtgagctgactcagaagcagccccctctgtgaacagctgtgtctg 420
 Db 361 TCTCTGTCTCTCTCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
 QY 421 cctccatcctcagacatcacaacagctgagcgtgagcgtctgagcgtctgagcgtctgagcgt 480
 Db 421 CCTCATCTCTCAGACATCACACCGTGTGAGCGTGTGAGCGTGTGAGCGTGTGAGCGT 480
 QY 481 tctcggcgaacacagcagagaccccgagcgggccccctcagaatctctgtgcatgtctggg 540
 Db 481 TCTCGCGCAACAGTGCAGAGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540

QY 541 gcttctcgtgctctctctccctgcccacacacagacatcctatgcatcaagttccactact 600
 Db 541 GCTTCTCGTGTCT 600
 QY 601 tccccaatgagctcctgtgtccagagtgccagacactgtacgtatcaagaacagctgtga 660
 Db 601 TCCCAATGAGTCCCTGTGTCCAGGTTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
 QY 661 tctacaattcattcattcagatcaacccctctccattctccctccctccctccctccctcc 720
 Db 661 TCTACAATTTCTATCTATCTAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
 QY 721 tcaatgtcctctactacatcctatggaatcagactcagaactaagaagaacactctctgagcag 780
 Db 721 TCAATGTCTCTCTACTACTCTATGCTGACACTGACACTGAAAGAAACAAATCTCTTGAGCAG 780
 QY 781 atgaaggaaatgcacaatattcaaaagacccctgcagaaatcagtcacaagaatgctgttg 840
 Db 781 ATGAAGGAATGCAAATATTCAAAAGACCTGACGAAATATGATCAACAGATGCTGTGTG 840
 QY 841 tcttgtctctagtgttctgtatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 900
 Db 841 TCTTGTCTCTAGT 900
 QY 901 gcttctgtgagagagtgagatggaatccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
 Db 901 GCTTTGTGAGAGAGTGAATGATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
 QY 961 caggctgtctctctctactcagctgagctgagctgagctgagctgagctgagctgagctg 1020
 Db 961 CAGGTGTCT 1020
 QY 1021 gccgctcagagcagcattccagaatgtgatctctctctcccaaaacagtgtgacccccc 1080
 Db 1021 GCCGCTTCAAGCGCAGCTCTCAGATGTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
 QY 1081 agcatgacccacagctgcacacctgcccagcaggaacactctccctgagaaatgacacttg 1140
 Db 1081 AGCATGACCCACAGCTTGCACACTTCCAGCGGAGCACTCTCTGACAGCAATGCCACTTGG 1140
 QY 1141 tgaagctgacccgaagatagatggtcccaattcccaatgcaagtcacacagcaaacctc 1200
 Db 1141 TGAAGCTGACCCGAAGATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
 QY 1201 accctcccaagacccctctcagtgagaagaatgtgaagaagaagaagaagaagaagaaga 1260
 Db 1201 ACCCTCCCAAGACCCCTCTCTAGTGAAGATGTGAAGATGTGAAGATGTGAAGATGTGA 1260
 QY 1261 ttaacaaacactgaatctctctcagagagctgactctcctc 1298
 Db 1261 TTAACAAACCTGAATCTCTTCAAGAGCTGACTCTCTCTC 1298

RESULT 2
 AX109234
 LOCUS 1658 bp DNA
 DEFINITION Sequence 1 from Patent WO0125269.
 ACCESSION AX109234
 VERSION AX109234.1 GI:13924107
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 1658)
 AUTHORS Deleersnijder, W., Berger, C., Loeken, C., Nys, G. and Venema, J.
 TITLE Human g-protein coupled receptor
 JOURNAL Patent: WO 0125269-A 1 12-APR-2001;
 Solvay Pharmaceuticals B.V. (NL)
 FEATURES
 source Location/Qualifiers
 1..1658
 /organism="Homo sapiens"

```

CDS
/db_xref="taxon:9606"
55..1302
/locus="IGS4 long version"
/codon_start=1
/db_xref="GI:13924108"
/translation="MSGMKLQNASWIIYQKLEDFQKHLNSTEYLAFLGPRRSHF
FLPVSVVPIFEVGVIGNVLVCLVLIQHOAMKPTNTYLFSLAVSDLLVLLGMPLE
VYEMRNYPLEPGVGYEKTALFEYTFASLITTVSVERVYAILHPRAKLOSTR
RRALRILOIVMGFSLPSLPTSTHIGIKFHPNLSLVPGSATCTVIRPMIYNITIQ
VTSLEFLPMTIVLSVLYLMALRLKDKSLKADGNANIQPKRSYKMLKLEVLV
FALCMAPHIDRLFEFSFEWSESLAANFVHVGVFLLSSAVNPIIYNLSRRF
QAFOMYISSFHKQMSQHPDLPAPQNRNIFLFEHLELIDIGPOPCOSSMNSH
LPTALSSQMSRTNYSFHNKT"

BASE COUNT      377 a      490 c      346 g      445 t
ORIGIN

Query Match      100.0%; Score 1298; DB 6; Length 1658;
Best Local Similarity 100.0%; Pred. No. 3,5e-259;
Matches 1298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 agggagagctcagccttgatcttaatgctcagggatggaataacttcagaatgcttcc 60
DB 29 AGGGAGAGCTCAGCCTTGATTTTATGTCAAGGATGAAAAACTTCAGAAATGCTTCT 88
OY 61 ggaatcaccagagaagaatcgaatcctcagaataaacttcagaacagccgaggaat 120
DB 89 GGATCTACAGAGAAACTGAAAGATCTTCAGAAACACCTGAAACACCGAGAGAT 148
OY 121 atctgagcctctctgaggaactcgcgagcgaactctctcccgctgctgctgctg 180
DB 149 ATCTGCGCTCTCTCTGCGGACCTCGCGGACGACACTTCTCTCCCGCTGCTGCTGT 208
OY 181 atgagcgaattctgctgctgctgctgctgctgctgctgctgctgctgctgctg 240
DB 209 ATGTGCAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 268
OY 241 agcaccagagctatggaagagcccaactactactctctcagcctgagcgtctcgac 300
DB 269 ACCACCAGGCTATGAAAGAGCCGACCACTACTACTCTTCAGCCGCGGCTCTGACC 328
OY 301 tccctgctcctgctctctggaatgagcctgaggtctatgagatgctgagcaactacc 360
DB 329 TCTTGCTCTGCTCTCTGGAATGCCCTTGAGAGTCTATAGATGTGCGCACTTACCT 388
OY 361 tctgctcggcgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 420
DB 389 TCTTGCTCTGCTCTCTGGAATGCCCTTGAGAGTCTATAGATGTGCGCACTTACCT 448
OY 421 cctcactcctggaatcacaacgctgagctgagctgagctgagctgagctgagctgag 480
DB 449 CCTTCATCTCTGAGCATCACACCGTGTAGGCTGAGCTGAGCTGAGCTGAGCTGAG 508
OY 481 tccgagcgaatactgagagcagccgagcgggagcctcctgagatccctcgagcatcgt 540
DB 509 TCCGCGCCCAAACTGCAAGACACCGCGCGGCTTCAGAGATCTTCGCGCATCTGTGG 568
OY 541 gctctcgcgtgctctctcctcctgccaacacagcaatcctgcaatcactact 600
DB 569 GCTTCTCCGCTCTCTCTCTGCGCAACACAGCATGCGATGCGATCAAGTTCCACTACT 628
OY 601 tccccaabtggtccctcggtcccaaggtctgagcactgtaagctgtaagcccaatgga 660
DB 629 TCCCCCAATGGGCTCGGTGCCAGGTTGCGCACCTGTACGGGTATCAAGCCCATGTGGA 688
OY 661 tctcaatctcactcagagctcctcctcctatctcactcctccccaatgagctgca 720
DB 689 TCTACATTTTCAATCAGAGCTCCTCTCTATTTCTACCTCTCCCATGACTGTCA 748
OY 721 tcaagtgcctcactaactcactgagcactaagaataagaacaactcttgaggcag 780
DB 749 TCAGTGTCTCTACTACTCATGTGCACTGAGACTAAAGAAAGCAATCTCTTGAGCAG 808

```

```

OY 781 atgaaggaatgcaatatatccaagaagccttgaggaataatcagtcacaagaatgcttg 840
DB 809 ATGAAGGAATGCAATATTTCAAGACCTCGCAAGAAATCAGTCAACAAAGATGCTTTGG 868
OY 841 tcttgctcagtgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 900
DB 869 TCTTGCTCTTAGTGTGCTATCTGTGGGCCCCGTTCCAGATGACCGACTCTTCTTCA 928
OY 901 gcttgtaggaagatgagtgaaatccctgctgctgctgctgctgctgctgctgctgctg 960
DB 929 GCTTGTGAGGAGGTGAGTGAATCCCTGCTGCTGTTCAACCTCGTCAATGTGTGTGT 988
OY 961 caagtgctctctcctcagagctgagctgctgctgctgctgctgctgctgctgctgct 1020
DB 989 CAGGTCTCTTCTCTACCTGAGCTCAGCTGCTCAACCCCATATCTATTAACCTATGCTGTC 1048
OY 1021 gccgctcagagcagatcaccagaatgctgctgctgctgctgctgctgctgctgctgct 1080
DB 1049 GCCGCTTCAGGAGCATTTCCAGAAATGTATCTCTTCTTCCCAAAACAGTGCGACTGCC 1108
OY 1081 agcatgacccacagcttgcaactgctgctgctgctgctgctgctgctgctgctgctg 1140
DB 1109 AGCATGACCCACAGTGTGCCACCTGCGGACGGAACATCTTCTGAGAGATGCGCACTTGG 1168
OY 1141 tggagctgagcgaagatagatgctgctgctgctgctgctgctgctgctgctgctg 1200
DB 1169 TGGAGCTGACCGAAGATATATGCTCCCAATTTCCATGTATCATGACACAACTGTC 1228
OY 1201 acctcccaagcgcctcctgagtgagacagatgcaagaacactcacaagcttccact 1260
DB 1229 ACTTCCCAACAGCCTCTCTAGTGAACAGATGTCAAGAACAACTATCAAGCTTCCACT 1288
OY 1261 ctacaacactgaaatctctcagagctgactcctc 1298
DB 1289 TTAACAACAACTGAATTTCTTCAAGACTGACTCTCTC 1326

RESULT 3
AX109236 1658 bp DNA PAT 30-APR-2001
LOCUS AX109236
DEFINITION Sequence 3 from Patent WO0125269.
ACCESSION AX109236
VERSION AX109236.1 GI:13924109
KEYWORDS
SOURCE
ORGANISM human.
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1658)
AUTHORS Deleersnyder, W., Berger, C., Loeken, C., Nys, G. and Venema, J.
TITLE Human g-protein coupled receptor
JOURNAL Patent: WO 0125269-A 3 12-Apr-2001;
Solway-Pharmaceuticals B.V. (NL)
FEATURES
source
1..1658
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
64..1302
/locus="IGS4 short version"
/codon_start=1
/db_xref="GI:13924110"
/translation="MEKLQNASWIIYQKLEDFQKHLNSTEYLAFLGPRRSHF
FLPVSVVPIFEVGVIGNVLVCLVLIQHOAMKPTNTYLFSLAVSDLLVLLGMPLE
VYEMRNYPLEPGVGYEKTALFEYTFASLITTVSVERVYAILHPRAKLOSTR
RRALRILOIVMGFSLPSLPTSTHIGIKFHPNLSLVPGSATCTVIRPMIYNITIQ
VTSLEFLPMTIVLSVLYLMALRLKDKSLKADGNANIQPKRSYKMLKLEVLV
FALCMAPHIDRLFEFSFEWSESLAANFVHVGVFLLSSAVNPIIYNLSRRF
QAFOMYISSFHKQMSQHPDLPAPQNRNIFLFEHLELIDIGPOPCOSSMNSH
LPTALSSQMSRTNYSFHNKT"

BASE COUNT      377 a      490 c      346 g      445 t
ORIGIN

```

Query Match 100.0%; Score 1298; DB 6; Length 1658;
 Best Local Similarity 100.0%; Pred. No. 3.5e-259;
 Matches 1298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 agggagagctcaggccttgatcttaatgctcaggatggaagaaacttcagaatgctcct 60
DB 29 AGGGAGAGCTCAGGCTTGATTTAATGTCAGGATGGAAGAACTTCAGAAATGCTTCT 88
QY 61 ggaatcaacagcagaactagaatccatctcagaagaacactgaaacagcaggagatc 120
DB 89 GGATCTACAGCAGAACTAGAAGATCCATTCAGAAACCCGAAACAGCAGCAGAGAT 148
QY 121 atctgacctctctcagcagcctcggcagcagcactctctcctcccgctgctggtg 180
DB 149 ATCTGGCCCTTCTCTGCGGACCTCGGCGCAGCCACTTCTTCTCCCGTGTGTGTGT 208
QY 181 atgtgcaaatcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 240
DB 209 ATGTGCAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 268
QY 241 agacacagcgtatagaagagcccaactactactctctcagcctcggcggtctctgac 300
DB 269 AGCACCAGGCTATGAGAGCGCCACCACTACTACTCTTACACCTGCGCGTCTGACC 328
QY 301 tctgtgctctgctctctgtgaatgccccctgagaggtctatgagatgtggcgaactac 360
DB 329 TCTGTGCTCTGCTCTTGTGAATGCCCTGAGAGTCTATGAGATGTGGCGCACTTACC 388
QY 361 tctgtgtggccgctggtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 420
DB 389 TCTTGTTCGGGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 448
QY 421 cctccatctctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480
DB 449 CCTCCATCTCAGCATCACCATCCGTCAGCGTGAAGCGTACGTCGTCATCTACCCGT 508
QY 481 tccgcagcaaaactgacagacacccggcgccggtccctcaggaatccctcggagatg 540
DB 509 TCCGGCCCAAACTGACAGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 568
QY 541 gctctcgtgtctctctctcctcctcctcctcctcctcctcctcctcctcctcctc 600
DB 569 GCTTCTCCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 628
QY 601 tccccaatgggtccctcgtgtcccgaggtgtggcagcctgctgagcagcagcagcag 660
DB 629 TCCCAATGGGTCCCTGGTCCAGGTTCGGCCACTGTACGGTATCATCAAGCCCATGT 688
QY 661 tctacaattctacatccagcagcagcagcagcagcagcagcagcagcagcagcagc 720
DB 689 TCTACAAATTTCAATCAGAGGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 748
QY 721 tcaagtctctctctctctctctctctctctctctctctctctctctctctctct 780
DB 749 TCAGTGTCTCTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 808
QY 781 atgaagaggaatgcaaatcttcaagaagccctgcaagaagaatcagcagcaagaatg 840
DB 809 ATGAAGGGAATGCAAAATTTCAAGAACCTTGCAAGAAATGAGCAACAGATCTGTT 868
QY 841 tcttggctcttaagtcttgcatactgtgtggcccgcttccacatctgacgagctctc 900
DB 869 TCTTGTGCTTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 928
QY 901 gcttctgtgagagagtgagatgaaatccctgtgctgtgtgtgtgtgtgtgtgtgt 960
DB 929 GCTTGTGAGAGAGTGTGAGATCCCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGT 988
QY 961 cagggtgtctctctctcagcagcagcagcagcagcagcagcagcagcagcagcag 1020
DB 989 CAGGTGTCTTCTTCTACTGAGCTCAGCTGTCAACCCCATTTATCTATTAACCTACT 1048
  
```

```

QY 1021 gcccttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1080
DB 1049 GCCGCTTCCAGGAGCATTCAGATGTGATCTCTTCTTCCCAACCAAGTGGCACTCC 1108
QY 1081 agcatgaccacagctgtgccacctgcccagcagacatctctcagagaatgacctgt 1140
DB 1109 AGCATGACCACAGTGTGCCACTGTCCAGCGGAACATCTTCTGACAGAAATGCCACT 1168
QY 1141 tggagctgacaggaatagatagtgctcccaatctcccatgtaagtcatactgaca 1200
DB 1169 TGGAGCTGACCGAAGATATAGTGTCCCAATTTCCATGTGATCATGCAACACTCT 1228
QY 1201 acctcccaagcagcctctcagtgaaacagatgcaagaacaaactcaagcttccact 1260
DB 1229 ACCTCCCAAGCGCCTCTCTAGTGAAAGATGTGCAAGAACAACTATCAAGCTTCC 1288
QY 1261 ttaacaaaactgaaatctctcagagctgaactctcctc 1298
DB 1289 TTAACAAAACCTGAATTTCTTTCAGAGCTGACTCTCTCC 1326

RESULT 4
AX109238 1658 bp DNA PAT 30-APR-2001
LOCUS AX109238
DEFINITION Sequence 5 from Patent WO0125269.
ACCESSION AX109238
VERSION AX109238.1 GI:13924111
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1658)
AUTHORS Deleersnijder, W., Berger, C., Loeken, C., Nys, G. and Venema, J.
TITLE Human g-protein coupled receptor
JOURNAL Patent: WO 0125269-A 5 12-APR-2001;
Solvay Pharmaceuticals B.V. (NL)
FEATURES
source Location/Qualifiers
1..1658
/organism="Homo sapiens"
/db_xref="taxon:9606"
55..1302
/note="IGS4B long version"
/protein_id="CAC37773.1"
/db_xref="GI:13924112"
/translacion="MSGKELQNASHTYQKLEDPQKHLNSTERYLAFLGCPRNSHF
FLPVSVVYPIFVVGIGNVLCVLVLOHQAARKTPNYLFSVLAVSDLVLLGMPLE
VYEMRNYPFLGPGVGCYKRTALFETVCPASILSTIVSERYVAIHLPRAKLQSTR
RALRLIGIVMGFSYSLPNTSINGIKHPYSPNGSLVSGATPKRKYIKPMYINFLIQ
VNSFLYFLPMTVISLYLALRLKDKHSPADSGMANIORPKRSYVKMFLVLYL
FAICNAPFHIDLFESFVEEMETLSLAAYFNVLHYVGYLFTYSSAVNPITLYMLLSRF
QAAFOVISFPHKQHSQHDPLPPAQRNIFLFECHFYELTEDIGPFLCGSSVNSH
LPTALSSSEDMSTNTQSFHFNKT"
BASE COUNT 378 a 489 c 346 g 445 t
ORIGIN

```

Query Match 99.5%; Score 1291.6; DB 6; Length 1658;
 Best Local Similarity 99.7%; Pred. No. 7.4e-258;
 Matches 1294; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 agggagagcctcaggccttgatcttaatgctcaggatggaagaaacttcagaatgctcct 60
DB 29 AGGGAGAGCTCAGGCTTGATTTAATGTCAGGATGGAAGAACTTCAGAAATGCTTCT 88
QY 61 ggaatcaacagcagaactagaatccatctcagaagaacactgaaacagcaggagatc 120
DB 89 GGATCTACAGCAGAACTAGAAGATCCATTCAGAAACCCGAAACAGCAGCAGAGAT 148
QY 121 atctgacctctctcagcagcctcggcagcagcagcagcagcagcagcagcagcagcag 180
  
```


D	b	149	ATATGCGCTTCCTCTGGCGACCTTCGGCGCAGCCACTTTCCTCCCGCTGTGTGGTGT	208
O	y	181	atgtgcgaattcttbtgttggggatcatatgtcaatgtccctgtgttcgttgcattctgc	240
D	b	209	ATGTGCAATTATTTGTGTGTGGGGGTCAATTGGCAATGTCCTGGTGTGGCGTGGATTCTGC	268
O	y	241	agcaaccaggtctatgtaaagacgcccacaactactactcttcagcctgtgcgtctctaac	300
D	b	269	AGCAACAGGGCTATGTAMACACCCACCAACACTACTCTTCCAGCCTGGGGGTCTGTGACC	328
O	y	301	tcctgtgttcctgtctcttggatgtccctctggaggtctctatgaatgttgggtgcaactac	360
D	b	329	TCCTGTGTCCTGTCTCTTGGAAATGCCCTGGAGGTCTATAGATGTGGCCCACTACCTT	388
O	y	361	tcctgttctgggcccgttgggtctgtacttcaagaacgcccctcttctagaacgtgtcttcg	420
D	b	389	TCCTTGTGTGGGCCGTGTGGCTGGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTGTCTTCG	448
O	y	421	ccctccatctctcaagcatcaccaacggtcaagcttggagcgtatcgttggccatctcacccgt	480
D	b	449	CCCTCATCTCTCAGCATCATCCACACCGTCAGCGTGAGGGCTACGGGACCATCTACACCGT	508
O	y	481	tcgcgcgcgaatctgtcabaagaccccggcgccgggcccctcaagatctcttggcatcgtctgg	540
D	b	509	TCCCGCGCCAACTGTCAAGACGACCCGGCGCCGCCCTCTCAAGATCCCTGGCACTCGTGG	568
O	y	541	gctctctcgtgtcctctctcctctgcacaaacacagatctcatgtatcgaattctcaact	600
D	b	569	GCTTCTCCGTCTCTTCTTCCCTGCGCAACACAGCATTCATGGCATCAATTCCACTACT	628
O	y	601	tcgcccaatgtgttcccttgttcccaaggtctcggccacctgtatcaggltcatcaagcccactgtga	660
D	b	629	TCCCAATAGGGTCCCTGCTGCCAGAGTTCGGCGCACCTGTACGGGTCACTAAGCCCATGTGGA	688
O	y	661	ctctcaatctcatcatcccaaggtcaactctctctctatcttcaactctctcccatagtcttca	720
D	b	689	TCTACAAATTTATCATATCCAGGTCACTCTCTCTTACTTACCTCCCTCCCATGACTGTCA	748
O	y	721	tcagatgtccctcttaactcctcatgtgcactcaagataagaagaacaaatctctcttggagag	780
D	b	749	TCACTGTCTCTACTACCTATGTGGCACTCAAGCTAAGAAGACAAATCTCTTGAAGCAG	808
O	y	781	atgaaaggaaatgcaaatatctcaaaagacccctgcagaataatcaatcaagaatgtcgttctg	840
D	b	809	ATGAAGGGAATGCMAATATTCAAAAGACCCGTGAGAAATACATCAACAAAGATGCTTTG	868
O	y	841	tcctgttctctagtgcttgcatactgttggcccgttcccaatgtacgcactctctctca	900
D	b	869	TCTTGTGCTTATGTTTGTTCATCTGTGTGGCCCCCGTCCACATTTGCACACTCTTCTTCA	928
O	y	901	gctttgttggaaaggatgtgaagtgaatccctctgtcgtctgttcaaacctgcgcatagtgagt	960
D	b	929	GCTTGTGTGAGAGACTGCATGAATCCCTGGCTGTCTGTTCAAACCTCGTCAATGTGTG	988
O	y	961	caagtgctctctcttctaaccttgagctcagctgttcaaccaccaatctataaactactgtctc	1020
D	b	989	CAGGTGCTTATTTCTACTGTAGGTCAAGCTGTCAACCCATTTATCTAATACACTACTCTC	1048
O	y	1021	gcgcgtctccaggtcagcatctccagaatgttatctctctcttccaaaacagtgtgaatccc	1080
D	b	1049	GCCCGTTTCCAGCGACGCAATTCACAAATGTGATCTCTTCTTCCACAAACAGTGTGGCACTCC	1108
O	y	1081	agcaatgaacccaagtctgcacacctgcgcagaaggagaactctctcttgcagaagaatgcacattg	1140
D	b	1109	AGCATGACCCCACTTGTCCACTGTGCCAGCGGAGAACATCTTCTCGAGAGATATCCACTTGTG	1168
O	y	1141	tgaggcttgaaccgaagataatagttccccaattctccatgtcatgtcatcatcgtacaactctc	1200
D	b	1169	TGGAGCTGACCGAAGATATAGGTGCCCAATTTCTATGTCAAGTATCCGTGCACAACTCTC	1228
O	y	1201	accttccaaacagccctctctagtgaacagatgttcaagaacaaactataaagcttcaact	1260
D	b	1229	ACCTTCCAAACGCCCTTCTATGTGAACAGATGTGTCAAGAACAAATCATCAAACTTCCACT	1288

QY	1261	ttacacaaacctgaattcttccagagctgactctcc	1298
Db	1289	TTAACAAACCTGAATCTTTCAGAGCTGACTCTC	1326
RESULT	5		
LOCUS	AX109240	1658 bp	DNA
DEFINITION	Sequence 7 from Patent WO0125269.		PAT
ACCESSION	AX109240		30-APR-2001
VERSION	AX109240.1	GI:13924113	
KEYWORDS	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1658)		
TITLE	Deleersnijder, W., Berger, C., Loeken, C., Nys, G. and Venema, J.		
JOURNAL	Human 9-protein coupled receptor Patent: WO 0125269-A 7 12-APR-2001; Solway Pharmaceuticals B.V. (NL)		
FEATURES	Location/Qualifiers		
source	1..1658		
CDS	/organism="Homo sapiens" /db_xref="taxon:9606" 64..1302 /note="IGS4B short version" /codon_start=1 /protein_id="CAC37774.1" /db_xref="GI:13924114" /translation="MEKLVNASMVIYQQLKEDPEQKHLNSTEYLAFICGPRSHFLP VSVNVPIFVGVIGNVLCVILIQAMKTPNLYLFLSIVSDLVLLGMLLEYE MNRNPLFEGPVGYEFTALFETVCFASLITVSEVERVALIHPRAKLOSTRRA LRILGIWGEFVLEFLPNTSINGIKFHPNSLVGSATCYIKPMIYNFLIOVTS FLFLYIPMTVILVLYLMAILRLKKKSLADGNAIOPCRKSNKMLFVLVLFAI CMARPHIDRLFESFVIEWETESLAAYNLVHYVGVLFYSSANPILYINLISRRFOA FQNVLSFRKQMSHQDPOLPPAQRNIFLIEGHFVELTIDIGHQFLCOSSVHNSHLP ALSSQMSSTNYQSFHFNKT"		
BASE COUNT	378 a	489 c	346 g 445 t
ORIGIN			
Query Match	99.5%;	Score 1291.6;	DB 6; Length 1658;
Best Local Similarity	99.7%;	Pred. No. 7.4e-258;	
Matches 1294; Conservative	0;	Mismatches 4;	Indels 0; Gaps 0;
QY	1	agggagagctcaagccttgatcttlaatgcaggaagaaacattcagaatgcttcc	60
Db	29	AGGGAGGCTCAGGCTTGATTTTAAATGACAGGATGGAANAATTCACAAATGCTTCT	88
QY	61	ggatctaccagcagaactagaaatccattccagaagaacacctgaacagcaggagat	120
Db	89	GGATCTACCAGCAGAACTAAGAATATCCATTTCCAGAAACACCTGAACAGCAGAGAGT	148
QY	121	atctggccttctctctcggaacctcgaggcgaagcaactctctccccggtctctggtgt	180
Db	149	ATCTGGCCTTCCCTCTCGGACCTCGGCGCAACACTTCTCTCCCGGTCTGTGTGCT	208
QY	181	atgtgcaaatctttgtgtaggggtcatgtgaaatgctcgtgtgtgctcgtgattctgc	240
Db	209	ATGTGCATTTTGTGTGTGGGGGTCAATGTGCATATGCTGTGTGCTGTGATTTTCC	268
QY	241	agcaccagatcagaagcagccaccacactactactcttcaagcttggcggctctgacc	300
Db	269	AGCACCAGGCTATGAAGACGCCACCAACTACCTCTTACACCTTGCGGTCTGTACC	328
QY	301	tctgtgtctgtctccttgaatgccctggaggtctatagatgttggcggaaactacctt	360
Db	329	TCTGTGCTCTGCTCTTGTGATGCCCTTGAGGCTATATGATGTGGCACAATACCTT	388
QY	361	tcttgtcggcccggtggtctgacttcaagaagcggccactcttgaagacggtgtgctcg	420

Db	389	TCITTTGGGGCCCTGGGCTGCTACTTTCAGACGGCCCTCTTTTGAAGACCGTGTGCTTCG	448
OY	421	ccctccatccctcaagca tcaaccaacgctcaagctgtagagcgctacgltggocaltctataccggt	480
Db	449	CCTTCATCCTCCTAGCATACACACCGTCAGGCTGGAGCGGTACTGTGGCAATCCTACACCGGT	508
OY	481	tcgcgagccaacactgcaagagcaaccggcgccgggcccctcaagatctctgggcatcgctctggg	540
Db	509	TCCGGGCCCAAACTGCAGAGACACCCGGCGCGGCGCTCTAGGATTCCTGGCAATCGTCTGGG	568
OY	541	gctctccgctgctctctcccccgcgccaacacagatacca tggcatcaagttccaaactact	600
Db	569	GCTTCTCCGCTGCTCTTCTTCCCTGGCCAAACACGATCATGTGGCATTCAGATTCTCACTACT	628
OY	601	tcgcccaatgggtgcccttggtctcccaaggtctcggcaacctgtacggtlca tcaagcccatgttga	660
Db	629	TCCCCAATGGGCGCCGCGGTCAGAGTGGCCACCGTACGGTCAATCAAGGCCCATGTGGA	688
OY	661	ctatacaattca tca tccaaggtcaacctctctctatcttaactctctcccatagtatgtlca	720
Db	689	TCTACAAATTCATCATATCCAGGTCACTCTCTTCTTATCACTCTCCCTCCCAATGACGTGA	748
OY	721	tcaagctctctca tcaacctca tgggaactcaagactcaagaagaacaatatctcttgaggag	780
Db	749	TCAAGTCTCTCTACTACTCACTGAGCACTGACTCAAGTAAGAAACAAATCTCTTGAAGGCKG	808
OY	781	atgaaggggaatgcaaatat tcaagaagaccttgagagaataatcag tcaacaaga tgcgctgtg	840
Db	809	ATGAAGGGGAATGCAAAATATATCAAGACCCCTGCAGAAATCACTCAACAAAGATGCTGTTGG	868
OY	841	tcttggtcttagtglttgctatctglttgggcccglttccaat tgaacgaactctcttca	900
Db	869	TCTTGTCTTAACTGATTTGCTATCTGTGTGGCCCGCTGCACATPAGACCACTCTTCTTCA	928
OY	901	gctttgtaggaagatggagtgatataccctbgygcgtgltgttaacctgfycaatlggytgt	960
Db	929	GCTTTGTGGAGAGTGGAGCTGATATCCCTGTGGCTGTGTGTTCAACCTCGTCATGTGTGTG	988
OY	961	caggtgctctctctcaacctgagctcaagctgtgtcaacoccca tta tctaataacctactgtctc	1020
Db	989	CAGGTGCTTATATTACCTGAGACCTCAAGCTGTCAACCCATTATCTATACTACTGTCTTC	1048
OY	1021	gccgcgtctccaggcagcat tccaagaatglatctctctcttccacaacaag tggcaactccc	1080
Db	1049	GCCGCTTCCAGGACACATCTCCAGAAATGTGATCTTCTTCCACAACAACAGTGGCACTCCC	1108
OY	1081	agcagagaccbaagrtgtccaactctgccagcggaabaatctctctgnaabagaatgcaacttgg	1140
Db	1109	AGCATGACCCCAAGTGGCCACTCTGCCACGGGAACATCTTCTGTGACAGATGCCACTTTTG	1168
OY	1141	tggagctgagcggagataatagttccccaattccca tgtacagcatccat tggacaactctc	1200
Db	1169	TGGACTGACCGAAGATATAGTGTCCCAATTTCTATGTAGTCAATCCGTGCACAACTCTC	1228
OY	1201	accctcccaacagccctctctctag tgaacaga tglcaagaacaaactat tcaaaagcttccact	1260
Db	1229	ACCTCCCAACAGCCCTCTCTACTGAAAGATGTCAAGAACAACATATCAAACTTCCAACT	1288
OY	1261	ttaacaaaaactgaaatctcttcaagagctgcaactctctc	1298
Db	1289	TTTAACAAAACCTGAATTTCTTTTAGAGGTGACATCTCTCTC	1326

RESULT	6				
AB041228					
LOCUS	AB041228	1248 bp	mRNA	PRI	21-SEP-2000
DEFINITION	Homo sapiens mRNA for G protein-coupled receptor TGR-1, complete cds.				
ACCESSION	AB041228				
VERSION	AB041228.1	GI:10257380			
KEYWORDS	G protein-coupled receptor TGR-1.				
SOURCE	Homo sapiens CDNA to mRNA.				
ORGANISM	Homo sapiens				

REFERENCE	Eumariyola: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
TITLE	Hosoya,M., Moriya,T., Kawamata,Y., Ohkubo,S., Fujii,R., Matsui,H., Shintani,Y., Fukusumi,S., Habeta,Y., Hinuma,S., Onda,H., Mishimura,O. and Fujino,M.
JOURNAL	Identification and Functional Characterization of a Novel Subtype of Neuromedin U Receptor
REFERENCE	J. Biol. Chem. 275 (38), 29528-29532 (2000)
AUTHORS	Shintani,Y., Moriya,T., Ohkubo,S. and Matsui,H.
JOURNAL	Direct Submission
TITLE	Submitted (03-APR-2000) to the DDBJ/EMBL/Genbank databases. Yasushi Shintani, Takeda Chemical Industries, LTD, Discovery Research Laboratories 1; 10 Madai, Tsukuba, Ibaraki 300-4293, Japan (E-mail:Shintani.Yasushi@takeda.co.jp, Tel.:81-298-64-5011, Fax:81-298-64-5000)
FEATURES	
source	Location/Qualifiers
gene	1..1248 /organism="Homo sapiens" /db_xref="taxon:9606"
CDS	10..1248 /gene="TGR-1" 10..1248 /gene="TGR-1" /codon_start=1 /product="G protein-coupled receptor TGR-1" /protein_id="BAB13721.1" /db_xref="GI:10257381" /translation="MEKLNQNSWITVYQQLDEPFQKHLNSTEEYLAFLLCGPRSHFFLPVSVAIPDFVVGVIGNVLCVLIOHQAMKTPNYLFSLASDLVALLLGMPLVEVELMRYPEFEGVCYFKTALPETGCFASISITVSERYVALIHPRAKIQSRRAA LRRIIGWGEVSFLSPNTSIHGIKFPKPPSGSLVPSASCTVTIKPMPIYNELIIONYS FLFLILPRTIVSYLYLMALRLPKDKLSLEADBNANQRCKSNKMPLVLYVFALI CWAPFHIDRLFEFTSEVESSELAAVENLVHVSVSFYLLSAVNPIIYNLSRFQA FQNNIFSHQMHSQHDPOLPEPADRNILECHPELVETLEDIGPOFPCOSSMHNSHLP ALSSQMSRTVYOSFHFNKT"
BASE COUNT	269 a 393 c 269 g 317 t
ORIGIN	
Query Match	96.1%; Score 1248; DB 9; Length 1248;
Best Local Similarity	100.0%; Pred. No. 8.3e-249;
Matches 1248; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Dn	27 atgtcaggatggaagaaccttcagaaagtcttccttgatctaccagcagaactagaagat 86 1 ATGTCAAGGATGGAAAACTTCAGAAATCCTTCTTGATCTACACAGCAAATAAGAAT 60
Oy	87 ccattccagaaaacaccttaaacgcgcgaagagatctcgacctctctctcggaactcgg 146 61 CCATTCCAGAAAACACTGAACAGCACGCCAGAGAGATGTGGCCCTTCTCGCGGACTCGG 120
Dd	147 cgcaagcaacttcctcccccgtgctctgtygtatgfygklatgfygcaaatlttgttggtaggg 206 121 CGCACCACTTCTTCTCCCGGTGCTGTCGTGTCGTATGTGGCCAATTTTGTGGTGGGGTTC 180
Oy	207 attggaactgtctcgtgtgtgtgcctgtgatctctgcaagcacaagactaagaagcgcac 266 181 ATTGGCAATGTCTCTGTGTGCCTGTGTGATTCGTGCAGCACACAGGCTATGAAGACGCCCCCACC 240
Oy	267 aactctacctcttagctgtgaggtctctgaacctccctggtctcctctcttggaatggcc 326 241 AACTACTACTCTTGTAGCGCTGGCGGTCTCTGACCTCTCGTGTCTCTTGGAAATGGCC 300
Oy	327 ctggaggtcatagatatgtgycgcaactaccccttctgtgtcggagcccgtyggctgtcac 386 301 CTGGAGGTCTATGAGATGTGGCGCAACTACCTTTCTTGTTCGGGCGCGTGGCGTCTAC 360
Oy	387 ttcaaagcagccctctttgaagacgtgtgtctgcgccatctctcaagcataccaacgctc 446 361 TTCAGAGCGCCCTTTTGAAGACCTGTGCTTCGGCTCATCTCAGCATCACCCCTTC 420

Db	541	AMACCAGCATCATGSGCATCAAGTTCCACTTTCCTCCAAATGGGTCCCTGGTCCAGGT	600
Qy	627	tcggccacccgtgaaggttcacaaagcccaatgtgaatcacaatltcaatltccaggtcac	686
Db	601	tcggccacccgtgaaggttcacaaagcccaatgtgaatcacaatltcaatltccaggtcac	660
Qy	687	tccttccctatcttaacctccctcccatagctgtcaatgtgtccctactaccitcatgtga	746
Db	661	tccttccctatcttaacctccctcccatagctgtcaatgtgtccctactaccitcatgtga	720
Qy	747	cccaagactaagaaagaaatctctcttgaggagaatgaggagaaatgaaatltcaaga	806
Db	721	ctcagactaagaaagaaatctctcttgaggagaatgaggagaaatgaaatltcaaga	780
Qy	807	ccctcagagaatctcagtcacaagaatgtctgtctgtctgtctgtctgtctgtctgt	866
Db	781	ccctcagagaatctcagtcacaagaatgtctgtctgtctgtctgtctgtctgtctgt	840
Qy	867	tcggcccccgttcacatgtacagcactctctctcagctctgttgagaagtgtgaatcc	926
Db	841	tcggcccccgttcacatgtacagcactctctctcagctctgttgagaagtgtgaatcc	900
Qy	927	ctggcgtcgtctgttcaacctgtgtcactgtgtgtgtgtgtgtgtgtgtgtgtgtgt	986
Db	901	ctggcgtcgtctgttcaacctgtgtcactgtgtgtgtgtgtgtgtgtgtgtgtgtgt	960
Qy	987	gctgtcaaacccatltcatataacctgtctcgcgcgtctccagcagcatltcagaat	1046
Db	961	gctgtcaaacccatltcatataacctgtctcgcgcgtctccagcagcatltcagaat	1020
Qy	1047	gtgactctctcttccacaagaatgtgacatccagcatgtacccacagttgtccacctgc	1106
Db	1021	gtgactctctcttccacaagaatgtgacatccagcatgtacccacagttgtccacctgc	1080
Qy	1107	caggcgaaatctctctctgtgacagaatgtgacactgtgtgtgtgtgtgtgtgtgtgt	1166
Db	1081	caggcgaaatctctctctgtgacagaatgtgacactgtgtgtgtgtgtgtgtgtgtgt	1140
Qy	1167	caattcccatgtcagcatcatcatgtcacaactctcaactctccaaagccctctgtgtga	1226
Db	1141	caattcccatgtcagcatcatcatgtcacaactctcaactctccaaagccctctgtgtga	1200
Qy	1227	csagatgtcaagaacaactatcacaagcttccacttcaacaaacctga	1274
Db	1201	csagatgtcaagaacaactatcacaagcttccacttcaacaaacctga	1248
RESULT	8		
AF242874	1239 bp	mRNA	PRI 13-JUL-2000
LOCUS			
DEFINITION	Homo sapiens neuromedin U receptor 2 (NMU2R) mRNA, complete cds.		
ACCESSION	AF242874		
VERSION	AF242874.1	GI:9082155	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 1239)		
AUTHORS	Howard,A.D., Wang,R., Pong,S.S., Mellin,T.N., Strack,A., Guan,X.M.,		
	Zeng,Z., Williams,D.L. Jr., Feighner,S.D., Nunes,C.N., Murphy,B.,		
	Stear,J.N., Yu,H., Jiang,Q., Clements,M.K., Tan,C.P., McKee,K.K.,		
	Hreniuk,D.L., McDonald,T.P., Lynch,K.R., Evans,J.F., Austin,C.P.,		
	Caskey,C.T., Van der Ploeg,L.H. and Liu,Q.		
TITLE	Identification of receptors for neuromedin U and its role in feeding		
JOURNAL	Nature 406 (6791), 70-74 (2000)		
MEDLINE	20351041		
REFERENCE	2 (bases 1 to 1239)		
AUTHORS	Liu,Q., McDonald,T.P., Wang,R., Jiang,Q. and Howard,A.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (09-Mar-2000) Pharmacology, Merck Research Labs, West		

FEATURES	POINT, PA 19486, USA
source	Location/Qualifiers
	1. .1239
gene	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="5"
	1. .1239
	/gene="NMU2R"
	1. .1239
	/gene="NMU2R"
	/note="FM-4"
	/codon_start=1
	/product="neuromedin U receptor 2"
	/protein_id="AA82735.1"
	/db_xref="GI:9082156"
	/translation="MEKIQNSMTVYQOKLEDPFQKHINSTSEYLAFLGPRSHFFLPVSVYVPEFVGVIGNVLYCIVLOHAMKPTNYLFEFLASDLVLVLMGPLEVE
	MRNRYPLFGFVGYCFPTALFETFCPSLISITTVSERYVALIHPRAKQSTRRA
	LRILGIWGFVLSFLPSTISLHGIKEHFEPGSLVPSATCTVIRKMTINFLVQTS
	FLFLPLTPVIVSLVLYLMLALRKDKSLEDEGNANQRPCKRSVNMKLVLYLFAI
	CMAPFIDRLFESEFVEEWSLEAIVNLIHVHVSQVFYLSAVNPILYNLSRFFQAA
	FQNYISSEHKHMSQHPOLPFAQRNIFLFECHVEVELTEDIGPQFPCOSSWNSHSLPTP
	ALSEQMSRTVQSFHFNK"
BASE COUNT	267 a 392 c 265 g 315 t
ORIGIN	
Query Match	95.5%; Score 1239; DB 9; Length 1239;
Best Local Similarity	100.0%; Pred. No. 6, 1e-247;
Matches 1239; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
36	atggaataactcagaatgctctcctggaatctaccagagaactagaatccattccag 95
1	ATGGAAAAAAGCTCAGAAATGCTCTTGATCTACAGAGAAACGTGAAGATCCATTCCAG 60
96	aaacacctgaacagacacgagagatctcgtctcctcctcgtcgaactcggcgagccac 155
61	AAACACCTGAACAGACGAGAGATCTGGGCTTCCTGGGACCTCGGCGAGCCAC 120
156	tctctccctcccccgtctcgttggttgatgttgcaaatctttgttggtgaggtcatgtgcaat 215
121	TTCTTCTCTCCCGCTCGTCTGTGTGTATGTGCCAATTTTGTGTGTGGGGTCAATTGGCAAT 180
216	gtccctgtgtgcctgtgattctcgaagacacgagctatgaaagacgcccacaactactac 275
181	GTCGTGTGTGTGCTGTGTATCTTCGACGACACGAGCTATGAAGAGCCACCAACTACTAC 240
276	cctctcagcctgtggcggtctctgaaactcctgtgtctcgtcctcttgaaatgcccctgtgaggtc 335
241	CTCTTCAGCTGGCGGTGTCTGACCTCTGTGCTCGTCTGCTCTTGGAATGCCCTGGAGGTC 300
336	tatgagatgtgacgaactaccctctctgtgttcggggccgttgaggtgtgtacttcaagaagc 395
301	TATGAGATGTGGCGCACTACCTCTTTTGTTCGGGGCCCGGTGGTCTACTTCAAGACG 360
396	ggcctccttgagaacggtgtgcttcgctcccaactcctcgaagacatcaacacgctcagcgttgag 455
361	GCCCTCTTTGAACCGGTGTGCTTCGCCATCCATCCACAGATCACCACCGTCAGCGTGGAG 420
456	cgctacgttgagcaatccataacccggttcggtcgccaaactgtcagaagacacccggcgccgggccc 515
421	CGCTACGTGGCCATCTACACCCGCTTCGGCGCAAACTGACAGACACCCGGGCGCGGGGCC 480
516	ctcgaagatccttgagatcgtctcgggggtcttcgcgtgctcttcctcctgtgccaacaacgagc 575
481	CTCAGAGATCTCTGGGATCGTCTGGGGCTTCTCCGTCTCTCTCCATCCCAACACCGACG 540
576	atcgaatggaatcaagttccaactctcccaacttggtgtccctggttccaggttcggaacac 635
541	ATCCATGGCATCAATTTCCACTACTTCCCAATGGGCTCCCTGTGTCACAGGTTCCGGCCACG 600
636	tgtaaagtcataacgccaatgtggaatctataaatctcatcatccaggttcaactcctcttca 695

```

Db 601 TGTACGGTCATCAAGCCCATGTGATCTACAAATTCATCATCCAGGTCACTCTTCTCTA 660
Qy 696 tttctactctctcccaatgagctgcaatgctctctactactactactgagacta 755
Db 661 tttctactctctcccaatgagctgcaatgctctctactactactactgagacta 720
Qy 756 aagaaagacaaatctctctgagcagatgaaagaaatgcaataatcaagacccctgaca 815
Db 721 AAGAAAGCAAAATCTTCTTACGAGCATGAGGAATGCAAAATATCAAGACCCCTGCACA 780
Qy 816 aaatcagtcacaaagatgctgttctgtctgtctgtctgtctgtctgtctgtctgtccg 875
Db 781 AAATCAGTCACAAAGATGCTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Qy 876 ttcacatgacgagactctctctacgactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 935
Db 841 tttccacattgacgagactctctctacgactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 900
Qy 936 gttgtcaacctgctcctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 995
Db 901 gttgtcaacctgctcctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
Qy 996 cccattatcataaactactgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1055
Db 961 cccattatctatgacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1020
Qy 1056 tcttctcaaaacagtgagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1115
Db 1021 tcttctcaaaacagtgagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1080
Qy 1116 atcttctacagaaatgacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1175
Db 1081 atcttctacagaaatgacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1140
Qy 1176 tgtcagtcacatgacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1235
Db 1141 tgtcagtcacatgacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1200
Qy 1236 agaaacaaatcatcaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1274
Db 1201 agaaacaaatcatcaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1239

RESULT 9
AX109242 1594 bp DNA PAT 30-APR-2001
LOCUS AX109242 9 from Patent WO0125269.
DEFINITION Sequence
ACCESSION AX109242
VERSION AX109242.1 GI:13924115
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1594)
AUTHORS Delorenz J, W., Berger, C., Loeken, C., Nys, G. and Venema, J.
TITLE Human g-protein coupled receptor
JOURNAL Patent: WO 0125269-A 9 12-APR-2001;
Solvay Pharmaceuticals B.V. (NL)
FEATURES
source 1..1594
/organism="Homo sapiens"
/db_xref="taxon:9606"
55..945
/note="IGS4A truncated DNA long version"
/codon_start=1
/protein_id="CAC37775.1"
/db_xref="gi:13924116"
/translation="MSGMRKLNASHYIOQKLEDPQKHUNTEEYLAFLCPRSHF
FLPVSVVYPIFVGVIGNVLCLVILQAKTKPNYLYSLVSLDVLKLPLE
VYEMRNYPFLPGVGCYFKTALFEVCPASILSTYVSEYVAILHFERKLOSTR
RRALRIIGVGFSLPNSLPGNSIKIKHYPNSLIVGSLVGSATGCVIKPMWYINFIQ
VTSFLYLLPMTVIVSYLYLMLRLKDKSLDADGNANIQPCRSKSVKMLSLMRSG

```

```

BASE COUNT 370 a 473 c 332 g 419 t
ORIGIN
VNPMLCSTSSMNCQVSSST"
Query Match 89.4%; Score 1160; DB 6; Length 1594;
Best Local Similarity 95.1%; Pred. No. 1.4e-230;
Matches 1234; Conservative 0; Mismatches 0; Indels 64; Gaps 1;
Qy 1 aaggagagctcagagccttgatattcaatgctcagagatggaataacttcaagatgtctct 60
Db 29 AGGGAGGCTCAGGCTTGATTTAAATGTCAGGATGGAAGAAAACCTTCAGAAATGCTTCT 88
Qy 61 ggaatcaccagcagaactagaagatccatcaagaacacacttgacagcagcagagat 120
Db 89 GGATCTACAGCAGAAACATAGAAAGATCCATTTCAGAAACCTTCAGAAACACCGAGAGACT 148
Qy 121 atctgactctctctgagacactcgagcagcacttctctcccgtytctgtgtgt 180
Db 149 ATCTGGCTTCTCTTGCGGAGCTCGGCGAGCAGCTTCTCCCGGTGTGTGTGT 208
Qy 181 atgtgcaatttttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 240
Db 209 ATGTGCAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 268
Qy 241 agcaccagactatgaagagccacacaaactactactcttcaagccttgagctgtgtgtgtgt 300
Db 269 AGCACCAGGCTATGAAAGAGCCGACCACTACTACTCTTTCAGCCTTGCGGTGTGTGTGTGT 328
Qy 301 tctgtgtctgtctcttctgtgaatgccccttgagagtgatctatgagatgtgtgtgtgtgtgt 360
Db 329 TCTGTGTCTGTCTCTTGTGAATGCCCTTGAGATGTATGATGTGTGTGTGTGTGTGTGTGT 388
Qy 361 tctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 420
Db 389 TCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 448
Qy 421 cctcaactcctcaagatcacacacgctgagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 480
Db 449 CCTCCATCTCTCAGCATCACACACGCTGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 508
Qy 481 tccgcgcacaaactgacagacacccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 540
Db 509 TCCGCCCAAACTGACAGACACCCGCGCGGCGCTTCAGATTCCTCGGCATCTGTGTGTGTGT 568
Qy 541 gcttctcggtgtctctctctcccgcccaacacagacatcagtgatcaagtctcactact 600
Db 569 GCTTCTCCGCTGCTCTCTCTCCGCGCCACACAGCATTCATGATGATGATGATGATGATGAT 628
Qy 601 tcccaatggtgtcccggtgtcccaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 660
Db 629 TCCCAATGGGTCCCTGTGTCAGAGTTCGCGCACTGTACGGGTATCAAGCCCATGTGTGA 688
Qy 721 tctacaattcatcacacagtgtaacctctctctctctctctctctctctctctctctctctctct 720
Db 689 TGTACAAATTTCAATCATCCAGGTGACTCTCTCTATTTTCACTCTCTCCCATGAGCTGTGA 748
Qy 721 tcaagtctctctactactactcctcatgagcactcaagactaaagaagaacaaatctcttgaggcag 780
Db 749 TCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 808
Qy 781 atgaagaggaatgcaaatatccaagaagccttgcaagaataatcagtcacaaagatgctgtttg 840
Db 809 ATGAAGGGAATGCAAAATATTAAGAGCCTTGAGAAATACAGTCAAGATGCTGTGTGTGTGT 865
Qy 841 tcttggtctttagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 900
Db 866 ----- 865
Qy 901 gcttctgagagagtgagtgaaatccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
Db 866 -CTTGTGTGAGAGGTGAGTGAATCCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 924

```


Db 1165 ACCTCCCAACAGCCCTCTCTAGTGAACAGATGTCAAGAACAACTATCAAGCTTCCACT 1224

Qy 1261 ttaacaaacctgaattcttcagagctgactctc 1298
|||||
Db 1225 TTACAAAACCTGAATCTTTTCAGAGCTGACTCTCCTC 1262

RESULT 11
AB041229
LOCUS Rattus norvegicus mRNA for G protein-coupled receptor TGR-1,
DEFINITION complete cds.
ACCESSION AB041229
VERSION AB041229.1 GI:10257382
KEYWORDS G protein-coupled receptor TGR-1.
SOURCE Rattus norvegicus cDNA to mRNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
Rattus.
1 (sites)
Moriya,T., Kawamata,Y., Ohkubo,S., Fujii,R., Matsui,H.,
Hosoya,M., Fukushima,S., Habata,Y., Hinuma,S., Onda,H.,
Shintani,Y., and Fujino,M.
Nishimura,O. and Fujino,M.
Identification and Functional Characterization of a Novel Subtype
of Neuromedin U Receptor
J. Biol. Chem. 275 (38), 29528-29532 (2000)
10887190
2 (bases 1 to 1314)
Fujii,R., Shintani,Y. and Hinuma,S.
Direct Submission
Submitted (03-APR-2000) to the DDBJ/EMBL/GenBank databases. Shuji
Hinuma, Takeda Chemical Industries, LTD, Discovery Research
Laboratories 1, Wadai 10, Tsukuba, Ibaraki 300-4293, Japan
(E-mail: Hinuma-Shuji@takeda.co.jp, Tel: 81-298-64-5035,
Fax: 81-298-64-5000)

FEATURES
source
1..1314
Location/Qualifiers
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
127..1314
/gene="TGR-1"
127..1314
/gene="TGR-1"
/codon_start=1
/product="G protein-coupled receptor TGR-1"
/protein_id="BAB13722.1"
/db_xref="GI:10257383"
/translation="MGKLENAWIMHPMKYLNSTEEYLAHLGPKRSDLSPLVAV
ALIFLVGVNLYCMIVIRHOTLKTPTNYLFSVLAVDLVLGMPLEIYEMHNY
PELFPGVCYFTALFETVCFASISLVTVSERVVAIVHPRAKLESTRRAFLIS
LWMSRVVPSLPTNSIHGIFQHPNGSSVPSASCTYKPKMYNLIQATSEFLYI
LPMRTISLYILMGIRLKRDSLENNKVAVNIHRSKSVTKMLFVLVPRICWTF
HVDRLFSEFVEWETSLAIVENLIHVGVEFYSSAVNPILYNIILSRFRAPNNV
SPTCKMCHPRHPOPOPPAKIIFLFECHLMELTEDAGPOFPQSSIHNTNLMAPCAG
EVP"

BASE COUNT 269 a 397 c 305 g 343 t

ORIGIN

Query Match 61.8%; Score 802.8; DB 10; Length 1314;
Best Local Similarity 78.8%; Pred. No. 1.6e-156;
Matches 976; Conservative 0; Mismatches 247; Indels 15; Gaps 1;

Qy 1 agggagagctcagggccttgatatttaatgtcagggatggaataacttcgaatgctcct 60
|||||
Db 92 AGGGACACCTGTGGTTGTGATTTTAAGCTCAGTAATGGAAACTTGAATAAGCTTCT 151
|||||

Qy 61 ggaactacagcagaagaacatagacatctcagaagaacacctgaacagcagcagagat 120
|||||
Db 152 GGATCCAC-----GATCCACTCATGTGAAGTACTTGAACACACAGAGAGAT 196
|||||

Qy 121 atctgagcttctctcgtgagacctcgcgacagcaacttctcctcccggtctgtgtgt 180
|||||

Db 197 ACTTGCCACCACCTGTGTGGAGCCCAACCGCAGTACCTATCCCTCCGGTGTCTGGACCT 256
|||||

Qy 181 agtgcacatttcttggtggtgggtgtaattggcaatgctctgtgtgctgtgattctgc 240
|||||

Db 257 ATGCGCTGATCTTCCATCGTGGGGGTAAATGGCAATCTTGTGTGTCATGATGATGTC 316
|||||

Qy 241 agaccagcctatgaagaagcaccacacactactactcttcagacctggagctctgacc 300
|||||

Db 317 GACATAGACTTTGAAGACACCACCACTACTATCTTTCAGCTTGGAGTCTCAGATC 376
|||||

Qy 301 tccctgctcctgctccttggaaatgccccttgaggatgtaataatgctgagcgaactac 360
|||||

Db 377 TGCTGTCTCTGCTCTGTGGGATGCCCTGGAATATCAGAGATGGGACCAATTAACCTT 436
|||||

Qy 361 tctgtctggcccgctgggctgctacttcaagaagcgccctcttggagcgtgtctcg 420
|||||

Db 437 TCCTGTTCGGGCTGTGGGATGCTACTTCAAGACACCCCTCTTCGAGACTGTGCTTTC 496
|||||

Qy 421 cctcacctctcagcatcacacacgctgagctgagagcgtctgagcctcctcaacccgt 480
|||||

Db 497 CCTCCATTCTCAGTGTCCACACGCTTAGCGTAAAGCGCTATGTGGCATTTGTCACCTT 556
|||||

Qy 481 tccgcgcaaacctgcaagacacccgagcggccctcagatccctcagatccctcgtctgg 540
|||||

Db 557 TCGAGCCCAAGGTGAGAGACGACGCGGCGACGGGCCCTCAGATTCCTCAGCTAGTCTGA 616
|||||

Qy 541 gcttctcgctctctctcctcctggccaacagaatcctatggcatcaatgtccactact 600
|||||

Db 617 GCTTCCTGTGGCTTTTCTTTGTCCTCAATACAGATCCATGGCATCAAGTTCACAGACT 676
|||||

Qy 601 tcccaatggtctcccgagcaggttgaggccacctgacgtctcctcagagcccatgtgag 660
|||||

Db 677 TTCACAGGGTCTCTCCGTAACCTGAGCTCAGCCACCTGCACTCACCAAAACCATGTGG 736
|||||

Qy 661 tctacaattctcatcaccagatccctctcctatcttcaactctccccaatgagctgta 720
|||||

Db 737 TGATTAATCTGATCATCCATCCAGCTACAGCTTCTCTTCAATCCCTCCAAATGACCTCA 796
|||||

Qy 721 tcaagtctcttaactaactcctcatgagcactaagaagaagaatcctctgagcag 780
|||||

Db 797 TCAGCCTCTCTACATACCTCATGAGGCTCAGGCTGAGAGATGATATCCCTTGGAGCGCA 856
|||||

Qy 781 atgaagaggaatgcaaatctcagaagccctgcaagaataatcagtaagaatgagcgttg 840
|||||

Db 857 ACAAACTGCTGTAATATTCACAGACCTCTAGAAAGTCACTACCAAGATGCTGTTTG 916
|||||

Qy 841 tcttggtcttaagtgtgtctatctgtgtggcccgcttccacactgacgaactctctca 900
|||||

Db 917 TCTGTGCTCTGCTGTTTGGCATCTGCTGACCCCTTCATGTGACCGGCTCTTCTTCA 976
|||||

Qy 901 gcttggtgaggaagtggagtgaaatccctgctgctgtgtgttcaactcgttcaatgtgtgt 960
|||||

Db 977 GCTTTCGTGAAGACGAGACAGAGTCCCTGCTGCTGTGTTCAACCTCATCATGCTGTAT 1036
|||||

Qy 961 cagggtctctcttcaactcagagctcagctgttcaacccaatctatcaactactgctc 1020
|||||

Db 1037 CAGGTCTCTTCTTATCTGACCTCGCGGTCAACCCATTAATTAATCACTCTGCTCTC 1096
|||||

Qy 1021 gccgcttccagcagcagatccagaaatgtaatctctcttctccacaacaagtgcaccc 1080
|||||

Db 1097 GCGCGCTTCGGGCGGCTTTCGAATAATGTGTCTCCCTCACTGCAAAATGGTCCATCC 1156
|||||

Qy 1081 agcatgacccaagtctgcaactctgccagcggaaacatctctctgacagaatgcaacttg 1140
|||||

Db 1157 GGCATCAGACAGGAGACTCTCAGCCCAAGATATCTTCTTGACAGATGTCACTCA 1216
|||||

Qy 1141 tgaagctgaacgaatataaggtctcccaattcccaatgtaagtcacatctgacaactctc 1200
|||||

Db 1217 TGGAGCTGACAGAGATGAGAGCCCCCAAGTCTCTGTGTCAGTCATTCATCCACACCA 1276
|||||

Qy 1201 acctccacaagcctctctatgtaacagatgttcaaga 1238
|||||

Db 1277 ACCTTACCATGCCCCCTGTGGCGAGAGGTACCATGA 1314

RESULT 12
LOCUS AF242875
DEFINITION Rattus norvegicus neuromedin u receptor 2 (NMU2R) mRNA, complete cds.
ACCESSION AF242875
VERSION AF242875.1 GI:9082157
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 1188)
AUTHORS Howard,A.D., Wang,R., Pong,S.S., Mellin,T.N., Strack,A., Guan,X.M., Zeng,Z., Williams,D.L., Jr., Feighner,S.D., Nunes,C.N., Murphy,B., Stalt,J.N., Yu,H., Jiang,Q., Clements,M.K., Tan,C.P., McKee,K.K., Hreniuk,D.L., McDonald,T.P., Lynch,K.R., Evans,J.F., Austin,C.P., Caskey,C.T., Van der Ploeg,L.H. and Liu,Q.
TITLE Identification of receptors for neuromedin U and its role in feeding
JOURNAL Nature 406 (6791), 70-74 (2000)
MEDLINE 20351041
REFERENCE 2 (bases 1 to 1188)
AUTHORS Liu,Q., McDonald,T.P., Wang,R., Jiang,Q. and Howard,A.D.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) Pharmacology, Merck Research Labs, West Point, PA 19486, USA

FEATURES
source
1..1188
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
1..1188
/gene="NMU2R"
1..1188
/note="NMU2R"
/codon_start=1
/product="neuromedin U receptor 2"
/protein_id="AF82756.1"
/db_xref="GI:9082158"
/translation="MGKLENASWIDHPLMKYLNSTREELAHLCGPKRSDLSLPVSAY ALFLVGVGCTFKALPEIVYCFASILSVITVSVERVAIVHPRAKLESTRRRLRLIS LWSFVSVPFSLPMTSIHGKFOHPNGSSVPSGATVTKPMVNYMLIQATSFELYL LPMTLISVLYLWGLRLKRDSELEANKVAVNIHRPSRKSYTKMLFVLVFAICMTPF HVDRLFESEWEWTESLAAVFNLIHVAGVEFLVSAVNPDIYNLSRRBRAFRNV SPKCKWCHPRHRQGRPRACKIIFLTECHLVELLEDAGRPOFGSSIHNNNTTAPCAAG EYP"

BASE COUNT 243 a 361 c 278 g 306 t

ORIGIN

Query Match 60.8%; Score 788.8; DB 10; Length 1188;
Best Local Similarity 79.4%; Pred. No. 1.3e-153;
Matches 953; Conservative 0; Mismatches 232; Indels 15; Gaps 1;

Db 1 ATGGGAAAACCTTGAATAATGCTCTCTGATCCAC-----GATCCACTCATG 45

Qy 36 atggaaaaactcagaatctctctctgatactcagcagcagaagaactagaagatcattccag 95

Db 1 ATGGGAAAACCTTGAATAATGCTCTCTGATCCAC-----GATCCACTCATG 45

Qy 96 aaacacctgaacacagcagcagagatcaltcgtcctctcctcgtcgagacctcgcgagccac 155

Db 46 AAGTACTTGAACACACAGAGAGTACTTGGCCACCTGTGTGAGACCAAGCGCAGTGAC 105

Qy 156 ttctctctcccggtctgtgtgtatgtatgtccaaattttgtgtgtgggtgtcatgtgcaat 215

Db 106 CTATTCCTTCCGCTGTGTGTGGCTATGCGCTGATCTTCTGTGTGGGGGTAATGAGCAAT 165

Qy 216 gtctctgtctcgtcgttgatctctcagcaccagcgtatgaagacgcccactactac 275

Db 166 CTCTGTGATGTCATGATGATGTCACATCAAGACTTTTGAAGACACCACCAACTACTACT 225

Qy 276 ctctcagcctgagcgtctctcagcctcctcgtctcctcctcgtgaatgcctcgtgagtc 335

Db 226 CTCTTACGCTTGGCACTTGTCCAGATGCTGTGCTCTCTCTGTGGGATGCTCTGGAAATC 285

Qy 336 tatgagatgtggtcgaac 395

Db 286 TAGGAGATGTGGCACAATTAACCTTCTCTTGTGGGCGTGTGGAGATGACTTCAAGACA 345

Qy 396 gccctcttgaacgctgtgtcgtcctccatcctcagatcaacacacacacacacacac 455

Db 346 GCCCTCTTGACAGCTGTGCTTGGCTTCATCTCAGTGTACCAACGGTTAGCGTTAGAG 405

Qy 456 cgtcagctgtgacatctcctacacccgtctccggtcccaactcagagacacccggtcggtc 515

Db 406 CGCTATGTGGCACTTGTCCACCTTTCGAGCCAGACCTGGAGACGCGGCGACGCGGC 465

Qy 516 ctcaagatcctcgacatcgtctgtgaggtctcctcgtcgtcctcctcctcgtcccaacac 575

Db 466 CTGAGGATCTTCAGCCTAGCTGTGAGCTTCTGTGTGTCTTCTTTCGCAATACACAC 525

Qy 576 atccatgtgacatcagatcctcctcctcccaatggttccctgtccaggttcggtccac 635

Db 526 ATCCATGTGACATCAAGTTCCAGCACTTTCACAGGGGTCCTGACTGACTGACTGACTG 585

Qy 636 tgaaggtacatcaagccacatgtgtatcaaatlcatcatcaggtcactcctccca 695

Db 586 TGCAAGTCCACCAACCAAGTGGGTATTAACCTTATCATCAACCTTACCACTTCCCTC 645

Qy 696 ttctacctcctcccatcagatcgtcctcagatgtcctcctcctcctcctcctcctcagatca 755

Db 646 TTCTCATCTCTCCCAATGACCTCATCAGGCTCTCTACTACTCATGAGGCTCAGAGCT 705

Qy 756 aagaagaacaaatcctcttgaagcagatgaaggaatgaataatcacaagccctgcaga 815

Db 706 AAGAGAGATGAATCCCTTGTGAGCGAACAAGTGGCTGTGAATTTTCACAGCCCTCTAGA 765

Qy 816 aaatcagtcacaagatgtcgttctgtctgtgtctatgttctcctcctcctcgttggcccg 875

Db 766 AAGTCAGTCCACCAAGATGCTGTGTGCTCTGTGCTCTGTGCTCTGTGCTCTGTGCTCT 825

Qy 876 ttccacatcagac 935

Db 826 TTCCATGTGACCGCGCTCTTCTTCTGAGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 885

Qy 936 gtgtcacaacctcgtcagatgtgtgtcaggtgtcctcctcctcagacacacacacacac 995

Db 886 GTGTCAACCTCATCCATGTGTGTATCATGAGTGTCTTCTTATCTGAGCTCCCGGTCAAC 945

Qy 996 cccattatctataactac 1055

Db 946 CCCATTATCTTAACCTCTGCTGCGGCTCTGCGGCGCTCTTGGAAATGTTGTCTCC 1005

Qy 1056 tcttccacaacaagatgtgac 1115

Db 1006 CTAACCTGCAAAATGATGTCATCCCGGATCGGCGCACAGGAGACTTCCAGCCCAAGATG 1065

Qy 1116 atcttctcgaagaatgcac 1175

Db 1066 ATCTTCTTGAAGATGTACCTGCTGTGAGCTGACAGAGATGTACAGGCCCGCCAGTCC 1125

Qy 1176 tctcagtcacatcagcacaactcactcactcctcctcctcctcctcctcctcctcctc 1235

Db 1126 GGTACAGTCATTCATCCACACACCAACTTACCAAGCGCCCTGTGTGAGAGAGGTACCA 1185

RESULT 13
AC008571/c
LOCUS AC008571 215088 bp DNA HTG
DEFINITION Homo sapiens chromosome 5 clone CTC-550M4, WORKING DRAFT SEQUENCE,

ACCESSION	4 unordered pieces.
VERSION	AC008571.4 GI:13699386
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE	1 (bases 1 to 215088)
AUTHORS	DOE Joint Genome Institute.
TITLE	Sequencing of Human Chromosome 5
JOURNAL	unpublished
REFERENCE	2 (bases 1 to 215088)
AUTHORS	DOE Joint Genome Institute.
TITLE	Direct Submission
JOURNAL	Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
	Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT	On Apr 20, 2001 this sequence version replaced gi:7708957.

```

Web site: http://www.jgl.doe.gov
-----
Project Information
Center Project Name: 396672, H361
Center clone name: CIT-HSPC_550M4

```

Summary Statistics

Consensus quality: 204117 bases at least Q40

Consensus quality: 210928 bases at least Q30

Consensus quality: 212668 bases at least Q20

Estimated insert size: 223000; pulse field gel estimation

Estimated insert size: 214788; sum-of-contigs estimation

Quality coverage: 6.63 in Q20 bases; pulse field gel estimation

Quality coverage: 6.89 in Q20 bases; sum-of-contigs estimation.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 4 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of 'N', but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

*	1	20385:	contlg of 20385 bp in length
*	20386	20485:	gap of unknown length
*	20486	35113:	contlg of 14658 bp in length
*	35144	35233:	gap of unknown length
*	35244	84714:	contlg of 49471 bp in length
*	84715	84841:	gap of unknown length
*	84815	215088:	contlg of 130274 bp in length

FEATURES	Location/Qualifiers
source	1. .215088

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clonene="CTC-550M4"
/clonene_lib="Caltech human BMC library C"
BASE COUNT 66410 a. 41474 c 42096 g 64807 t 301 other
ORIGIN

```

Query Match	58.0%	Score 752.8	DB 2	Length 215088
Best Local Similarity	99.7%	Pred. No. 2.3e-146		
Matches 754	Conservative	0	Mismatches 2	Indels 0
				Gaps 0

Oy	1	aggagagagcctcagagcccttgcgaattcctaagtcagagatcgagaaacattccagaatgcttccct	60
Db	105587	AGGGAGGCTCAGGCCCTTGATTTAATGTCAGGAGATGAAAAACTTCACAAATGCTTCCCT	105528
Oy	61	ggatctaccagcagaaactagaaagatccattccagaaacacctgaaacagcaccaggaagt	120
Db	105527	GGATTTCACACACAGAAATTCAGAAATCATCTTCAGAAACACCTTAACAGCAGCGAGAGT	105468

Oy	121	atcggccttcctctctgaggactcggcgagccactcttcctcccgctcgtcgtgtc	180
Db	105467	ATCTGGCCCTTCCTCGCGAGACTCTGGGGCACCCACTTCTTCGCCGTCGTGTGTGT	105408
Oy	181	atgtgccaattcttgtagtgggtcaattgacaatgtccggtagtgccgtgatctgc	240
Db	105407	ATGTGCCAATTTTGTGTGGTGGGGTCAATTGGCAATGTCTGTGGTGTGCTTGTGTC	105348
Oy	241	agcacagagctatgaagacgcccacactactactctctcaagcctctgcygctctgacc	300
Db	105347	AGCACAGAGCTATGAGAGCGCCACACAATCTACCTCTTTCAGCGTGGCGTGTGAC	105288
Oy	301	tctcgtctcgcctcctctggaaatgcctctgaggtctctatgagatgttgcgcaactccct	360
Db	105287	TCTCTGGTCTCTCTCCCTTGGAAATGCCCTGTGAGGTCTATGAGATGTGGCGCACTACCC	105228
Oy	361	tctgtctcgggcccgttgagctgcctacttcaagaacgcccctcttgagaccgttgtctcg	420
Db	105227	TCTTGTTCGGGGCCGTGGGCTGCTACTTTCAGAGCGGCCCTTTTGAACCGTGTGCTTCG	105168
Oy	421	ccctcatcctcagatcacaccacgctcagcgttgagacgctacgttgcacatcctacaccgt	480
Db	105167	CCCTCATCTCTCAGCATCACACCGTCAAGCGTGGAGCGCTAGTGTGGCCATCTACACCG	105108
Oy	481	tccgcgcgaacactgcagagaccccgcgcgccctccagatctctgcgcatcgtctcgg	540
Db	105107	TCCGCGCCCAACTGCAGAGCACCCGGGGCCGGGCTCTCAGATCTCTTGGCATCGTCTGGG	105048
Oy	541	gctctcgcgtcctctctccctccgcacacacagcatcatctgagataagtctccact	600
Db	105047	GCTTCTCGGCTCTCTTCTCCCTGCGCCACACACACATCATGTCATAGTTCCACTACT	104988
Oy	601	tcoccaatgagctcccttggtcccaaggtctcgcgacactctgaagtcataagcccactgtga	660
Db	104987	TCGCCAATGGGTCCCTGTGCTCCAGAGTTCGGGCACCTGTAGCGTATCAAGCCATGTGGA	104928
Oy	661	tctacaattcatcatccaggctcaactctctctatcttactctctccctcccatgactgtca	720
Db	104927	TCTACAAATTCATCATCCAGGTCAACCTCTTCTATCTCACTCTCCCATGACTGTCA	104868
Oy	721	tcagtgctctactactcaatgacatcaagaa 756	
Db	104867	TCAGTGTCCTCTACTACTCATGCGACTCACTCAAGTGA 104832	
RESULT 14			
AX147772/c	AX147772	801 bp	DNA
LOCUS	AX147772	Sequence 17 from Patent WO0136473.	PAT
DEFINITION	AX147772		08-JUN-2001
ACCESSION	AX147772		
VERSION	AX147772.1	GI:14346807	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
	1 (bases 1 to 801)		
	Vogeli,G., Wood,L.S., Parodi,L.A., Hiesch,R.R., Lind,P.,		
	Slightom,T., Schellin,K.A., Kayes,P.S., Bannigan,C.M., Ruff,V.,		
	Sejltz,T. and Huff,R.M.		
	Novel g protein-coupled receptors		
	Patent: WO 0136473-A 17 25-MAY-2001;		
	PHARMACIA & UPJOHN COMPANY (US)		
FEATURES	Location/Qualifiers		
SOURCE	1..801		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	200 a 187 c 261 g 153 t		
ORIGIN			
Query Match	55.88;	Score 723.8;	DB 6;
Best Local Similarity	99.7%;	Pred No. 4;	le-140;

Matches 725: Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
QY	30	tcagggatggaaaaacttcagaatgcttcctggatctacacagcagaacacagaatcca	89	
Db	801	TCAGGGATGGAAAAACTTCAGAAATGCTTCTGGATCTACACAGAAACTAGAAAGATCCA	742	
QY	90	ttccagaagaacacttgacagacacagagaglatctgacctctctctgcygacctcggc	149	
Db	741	TTCCAGAAACACCTGAAACACACAGAGAGATATCTGSCCTTCCCTCGGACCTCGGCGC	682	
QY	150	agcgaacttctctcccccggctgctgctgctgctgctgctgctgctgctgctgctgctg	209	
Db	681	AGCAGACTTCTTCCCTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	622	
QY	210	ggcaatgctcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg	269	
Db	621	GGCAATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	562	
QY	270	taactaaccttcagccttgccgctcctgacccctgctcctgctccttggaaatggccctg	329	
Db	561	TACTACCTTTACAGCCCTGGGGCTCTGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	502	
QY	330	gaagctcatagatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg	389	
Db	501	GAGGTCATAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	442	
QY	390	aagagccctctcttgagacgctgctgctgctgctgctgctgctgctgctgctgctgctg	449	
Db	441	AAGAGCCCTCTTTGAGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	382	
QY	450	gtggagcctcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg	509	
Db	381	GTGGAGCGCTACGCTGCGCATCTACACCGCTGCGCGCAAGCTGAGAGACACCGCGCG	322	
QY	510	cgggcccctcagatccctcggcctgctgctgctgctgctgctgctgctgctgctgctg	569	
Db	321	CGGGCCCTCAGGATCCCTGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	262	
QY	570	accagcatcagatgcatcagatgctcactatctcccaatggctccctgctcagagctg	629	
Db	261	ACCAAGCATCATGCGATCAAGTTCACACTACTTCCCAATGGTCCCTGGTCCCAAGTTCG	202	
QY	630	gccacctgtaagctcacaagcccaatgctgctgctgctgctgctgctgctgctgctg	689	
Db	201	GCCACCTGATGCTCATCAAGCCCATGTGATCATCAATTTCAATCATCAGTTCACCTCC	142	
QY	690	ttcctatctaacctcccccacatgactgctgctgctgctgctgctgctgctgctgctg	749	
Db	141	TTCTATTTTACCTCTCCCTCCCATGACTGTATCATAGTCTCTACTACTATGCTGCTG	82	
QY	750	agactaa 756		
Db	81	AGAGTGA 75		
RESULT	15			
AX139107	AX139107	729 bp DNA	PAT	30-MAY-2001
LOCUS	Sequence 1	from Patent EP1090990.		
DEFINITION	AX139107			
ACCESSION	AX139107			
VERSION	AX139107.1	GI:14274786		
KEYWORDS				
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 729)			
AUTHORS	Harland, L.P.			
TITLE	Human g-protein-coupled receptor			
JOURNAL	Patent: EP 1090990-A 1 11-APR-2001;			
	Pfizer Limited (GB) ; PFIZER INC. (US)			
FEATURES	Location/Qualifiers			
	1..729			

BASE COUNT 135 a 245 c 167 g 182 t				
ORIGIN				
/organism="Homo sapiens"				
/db_xref="taxon:9606"				
Query Match 55.3%; Score 717.8; DB 6; Length 729;				
Best Local Similarity 99.7%; Pred. No. 7.3e-139;				
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
QY	36	atggaaaaaacttcagaatgcttcctggatctacacagcagaacactagaatcatccag	95	
Db	1	ATGGAAAAACTTCAGATGCTTCTGATCTACAGCAGAACTAGAAAGATTCATTCAG	60	
QY	96	aaacacctgaagcagaccggagagatctgacctctctcctgagacctggcgcaaccac	155	
Db	61	AAACACCTGAACAGCACCGAGAGTATCGCTTCTCTGCGGACCTCGGCGGACGAC	120	
QY	156	ttctccctcccgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg	215	
Db	121	TTCTTCTCTCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	180	
QY	216	gtcctggtgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg	275	
Db	181	GTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	240	
QY	276	ctcttcagccttgccgctgctgctgctgctgctgctgctgctgctgctgctgctgctg	335	
Db	241	CTCTTACGCTTGCGGGCTGTGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	300	
QY	336	latgagatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg	395	
Db	301	TATGAGATGCTGCGCAACTACCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	360	
QY	396	gccctcttgagacgctgctgctgctgctgctgctgctgctgctgctgctgctgctg	455	
Db	361	GCCTCTTTTGAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	420	
QY	456	cgtacgtgacatctcacaacccgctccggcgaacactgagagacccggcgccgggac	515	
Db	421	CGCTACCTGGCATCTTACACCGCTTCCGCGCAAACTGACAGCACCGCGCGCGGCG	480	
QY	516	ctcagatctcggcatcgtctggtgcttccgctgctcctcctcctgcccacaacacag	575	
Db	481	CTCAGATCTCTGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	540	
QY	576	atccatggcatcaagtccaactctcccaatgggtccctgctcccaagcttgcgcac	635	
Db	541	ATCCATGGCATCAAGTTCACCTTCTCCCAATGGGTCCCTGCTCCAGGTTGGGCACC	600	
QY	636	tgtaagctcaagcccatgctgctgctgctgctgctgctgctgctgctgctgctgctg	695	
Db	601	TGTACGCTCATCAAGCCCATGTGATCTACAAATTTATCATTCAGGTCACTCTTCTTA	660	
QY	696	ttctaacctctcccaatgactgctgctgctgctgctgctgctgctgctgctgctgctg	755	
Db	661	TTCTTACCTCTCTCCCATGATGATCATGATGCTCTCTACTACTCATGAGACTCAAGTG	720	
QY	756	a 756		
Db	721	A 721		

Search completed: April 22, 2002, 16:20:16
Job time: 10975 sec

THIS PAGE BLANK (USPTO)

PF 13-DEC-2000: 2000MO-US33787.
 XX
 PR 17-DEC-1999: 99US-0466435.
 PR 25-APR-2000: 2000US-0558099.
 PR 30-JUN-2000: 2000US-0609146.
 XX
 PA (SYNA-) SYNAPTIC PHARM CORP.
 PI Bonini JA, Lerman GS, Quan Y, Ogozalek K;
 XX
 DR WPI, 2001-390240/41.
 DR P-PSDB; AAE03629.
 XX
 PT A purified mammalian SNORF62 or SNORF72 receptor protein for
 PT identification of compounds to treat e.g. inflammation, arthritis,
 PT autoimmune diseases, transplant rejection, AIDS, cancer -
 PS
 PS Claim 42: Fig 3; 256pp; English.

XX The invention relates to human G-protein coupled receptors, SNORF62 and
 CC SNORF72 and their corresponding cDNA molecules. SNORF62 and SNORF72
 CC receptors are specific for neuromedin U (NMU) neuropeptides, hence they
 CC are also known as NMU receptors. The agonist and antagonist of NMU
 CC receptors are useful for treating an abnormality in a subject that is
 CC alleviated by decreasing or increasing the activity of NMU receptor.
 CC The NMU receptors serves as a valuable tool for designing drugs which are
 CC useful for treating various pathophysiological conditions such as
 CC inflammation, arthritis, autoimmune diseases, transplant rejection,
 CC graft vs host disease, bacterial, fungal, protozoan and viral infections,
 CC septicemia, AIDS, pain, psychotic and neurological disorders, including
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, neuromotor disorders, respiratory disorders, asthma,
 CC eating/body weight disorders including obesity, bulimia, diabetes,
 CC anorexia, nausea, hypertension, hypotension, vascular and cardiovascular
 CC disorders, ischemia, stroke, cancers, sexual disorders, circadian
 CC disorders, renal disorders, bone diseases including osteoporosis, benign
 CC prostatic hypertrophy, gastrointestinal disorders, nasal congestion,
 CC dermatological disorders such as psoriasis, allergies, Parkinson's
 CC disease, Alzheimer's disease, acute heart failure, angina disorders,
 CC delirium and dyskinesias such as Huntington's disease. They can also be
 CC used to regulate steroid hormone disorders, epinephrine release
 CC disorders, electrolyte balance disorders, endocrine disorders, memory
 CC disorders, somatosensory disorders, metabolic disorders, behavioural
 CC disorders, drug addiction, migraine, Addison's disease, Cushing's
 CC disease, prevent miscarriage, induce labour or to treat dysmenorrhoea.
 CC The present cDNA sequence encodes human G-protein coupled receptor,
 CC SNORF72.
 CC
 XX
 SO Sequence 1298 BP; 278 A; 405 C; 282 G; 333 T; 0 other;

Query Match 100.0%; Score 1298; DB 22; Length 1298;
 Best Local Similarity 100.0%; Pred. No. 2, 3e-301;
 Matches 1298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agggaggagctcagagcttgatcttaatgcagggatggaagaaactcagaatgctcct 60
 DB 1 agggaggagctcagagcttgatcttaatgcagggatggaagaaactcagaatgctcct 60
 QY 61 gggcttcacagcagaactcagaatcattccagaaacccggaagacccgagagagt 120
 DB 61 gggcttcacagcagaactcagaatcattccagaaacccggaagacccgagagagt 120
 QY 121 atctgagctcctctcgcagacctcgcgcagccacctctccctcccggtctcgtggtc 180
 DB 121 atctgagctcctctcgcagacctcgcgcagccacctctccctcccggtctcgtggtc 180
 QY 181 atctgagcaattcttggtggtggggtcattcggcaatgctccgtgtgctcgtgattctgc 240
 DB 181 atctgagcaattcttggtggtggggtcattcggcaatgctccgtgtgctcgtgattctgc 240
 QY 241 agcacccggctatgaaagcccccacactactaccctctcagcctggcggtctctgacc 300
 DB 241 agcacccggctatgaaagcccccacactactaccctctcagcctggcggtctctgacc 300

DB 241 agcacccggctatgaaagcccccacactactaccctctcagcctggcggtctctgacc 300
 QY 301 tctctgagctcgtcctcttggaatgccccctggaagttctatgagatggtgcgaactaccct 360
 DB 301 tctctgagctcgtcctcttggaatgccccctggaagttctatgagatggtgcgaactaccct 360
 QY 361 tctctgagccgctggtgctgacttcaagacgagccctcttgagacggtgctctg 420
 DB 361 tctctgagccgctggtgctgacttcaagacgagccctcttgagacggtgctctg 420
 QY 421 cctccatcctcagaatcacaccgctcagctgagagctagctgtgccaatccatccaccgt 480
 DB 421 cctccatcctcagaatcacaccgctcagctgagagctagctgtgccaatccatccaccgt 480
 QY 481 tccggagcaatcagcagagacccggcgccgcccctcagatctcccgatgctgtgg 540
 DB 481 tccggagcaatcagcagagacccggcgccgcccctcagatctcccgatgctgtgg 540
 QY 541 gctctcgtcgtctctctccctgcccacacacagatccatggtcagatccactact 600
 DB 541 gctctcgtcgtctctctccctgcccacacacagatccatggtcagatccactact 600
 QY 601 tccccaatggtcccggtgctccaggttcggccacctgacgtcaccatcaagccatgga 660
 DB 601 tccccaatggtcccggtgctccaggttcggccacctgacgtcaccatcaagccatgga 660
 QY 661 tctcaaatccatcatccaggtacacctctctatctccctccctccccaatgagctga 720
 DB 661 tctcaaatccatcatccaggtacacctctctatctccctccctccccaatgagctga 720
 QY 721 tcaagtctcctctactactcactcactcactcactcactcactcactcactcactcact 780
 DB 721 tcaagtctcctctactactcactcactcactcactcactcactcactcactcactcact 780
 QY 781 atgaagggaatgaaatataatcaagaagccctcgagaatataatgacaaagatgctgtt 840
 DB 781 atgaagggaatgaaatataatcaagaagccctcgagaatataatgacaaagatgctgtt 840
 QY 841 tcttgagctcaggtgtgctcactgtgtggcccggtccacatctgacgagctctctca 900
 DB 841 tcttgagctcaggtgtgctcactgtgtggcccggtccacatctgacgagctctctca 900
 QY 901 gctttgtgagaggtgagtgaaatccctggtcgtgtgttcaacctgtcatgtgtgt 960
 DB 901 gctttgtgagaggtgagtgaaatccctggtcgtgtgttcaacctgtcatgtgtgt 960
 QY 961 caagtgtcctctctcactcactcactcactcactcactcactcactcactcactcact 1020
 DB 961 caagtgtcctctctcactcactcactcactcactcactcactcactcactcactcact 1020
 QY 1021 gccgcttcagagcagatcccaaatgtgactctctctctctcacaacaacagtgagctccc 1080
 DB 1021 gccgcttcagagcagatcccaaatgtgactctctctctctcacaacaacagtgagctccc 1080
 QY 1081 agcatgacccaagttgcccacccgccaagcagacaatctcccgacagaaatgccaattg 1140
 DB 1081 agcatgacccaagttgcccacccgccaagcagacaatctcccgacagaaatgccaattg 1140
 QY 1141 tggagctgacaggaatagatggtccccaatctcccatgtaagtaagtaatcacaactctc 1200
 DB 1141 tggagctgacaggaatagatggtccccaatctcccatgtaagtaagtaatcacaactctc 1200
 QY 1201 acctcccaaacagccctctcagtgaaacagatgtcacaagaacaaactatcaaaagctccact 1260
 DB 1201 acctcccaaacagccctctcagtgaaacagatgtcacaagaacaaactatcaaaagctccact 1260
 QY 1261 ttaacaaaactgaaattcttcaagagctgactctc 1298
 DB 1261 ttaacaaaactgaaattcttcaagagctgactctc 1298

RESULT 2
 AAF80322

ID AAF80322 standard; DNA: 1658 BP.
XX AAF80322:
AC
XX 29-JUN-2001 (first entry)
DT
XX
DE Nucleotide sequence of a G-protein coupled receptor IGS4A long version.
XX
KM Human: G-protein coupled receptor: IGS4; IGS4B; schizophrenia;
KM nervous system disorder; psychiatric disorder; Parkinson's disease;
KM episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KM cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KM motility disorder; myocardial infarction; hypertension; dyslipidemia;
KM gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KM inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KM gynecological disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 55..1302
FT /tag= a
FT /product= "G-protein coupled receptor IGS4A long version"
XX
PN WO200125269-A2.
XX
XX 12-APR-2001.
XX
PF 25-SEP-2000; 2000WO-EP09584.
XX
XX 24-SEP-1999; 99EP-0203140.
PR 24-SEP-1999; 99NL-1013140.
PR 28-JUL-2000; 2000EP-0202683.
PR 31-JUL-2000; 2000US-0222047.
XX
XX (SOLV) SOLVAY PHARM BV.
PA
XX Deleernijder W, Berger C, Loeken C, Nys G, Venema J;
XX
XX WPI: 2001-273568/28.
DR P-PSDB: AAB67802.
XX
XX New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers
XX
XX Claim 1: Page 75-77; 102pp; English.
PS
XX The present sequence encodes the long version of a human G-protein
CC coupled receptor designated IGS4A. IGS4 exists in two polymorphic forms,
CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
CC for preventing, ameliorating or correcting dysfunctions or diseases.
CC These diseases include peripheral nervous system, psychiatric and central
CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
XX Sequence 1658 BP; 377 A; 490 C; 346 G; 445 T; 0 other;

Query Match

100.0%; Score 1298; DB 22; Length 1656;

Best Local Similarity 100.0%; Pred. No. 2,5e-301;
Matches 1298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 aggggaagctcaagccttgatatttaatgctcaaggaatgtaaaaaactcagaatgctcct 60
DB 29 aggggaagctcaagccttgatatttaatgctcaaggaatgtaaaaaactcagaatgctcct 88
QY 61 ggaatcacagcagaactgaagaatccatccagaacacccggaacagcagaagagt 120
DB 89 ggaatcacagcagaactgaagaatccatccagaacacccggaacagcagaagagt 148
QY 121 atcgtgctccctctcgagaccctcgcgacagcaactctctcccccgtctcgtgtgtg 180
DB 149 atcgtgctccctctcgagaccctcgcgacagcaactctctcccccgtctcgtgtgtg 208
QY 181 atgtgccaatctctgt 240
DB 209 atgtgccaatctctgt 268
QY 241 agcacagcgtatgaagagcaccacacacacacacacacacacacacacacacacacac 300
DB 269 agcacagcgtatgaagagcaccacacacacacacacacacacacacacacacacacac 328
QY 301 tccctgtccctcctctgtggaatgccccctgagaggtctatgaatgtgtgtgtgtgtgt 360
DB 329 tccctgtccctcctctgtggaatgccccctgagaggtctatgaatgtgtgtgtgtgtgt 388
QY 361 tcttgt 420
DB 389 tcttgt 448
QY 421 cctccac 480
DB 449 cctccac 508
QY 481 tccgcgcgaac 540
DB 509 tccgcgcgaac 568
QY 541 gcttctcgt 600
DB 569 gcttctcgt 628
QY 601 tccccaatgt 660
DB 629 tccccaatgt 688
QY 661 tctacaattcatcatcagatcagctcctctctatctcctctcctcccaatgactgtca 720
DB 689 tctacaattcatcatcagatcagctcctctctatctcctctcctcccaatgactgtca 748
QY 721 tcaagtctcctctactactactactactactactactactactactactactactactact 780
DB 749 tcaagtctcctctactactactactactactactactactactactactactactactact 808
QY 781 ataaaggaatgtcaaatatcaagaacccctcgagaaatcaatgaatgaatgtgtgtgt 840
DB 809 ataaaggaatgtcaaatatcaagaacccctcgagaaatcaatgaatgaatgtgtgtgt 868
QY 841 tcttggtctagt 900
DB 869 tcttggtctagt 928
QY 901 gctttgtgagaggt 960
DB 929 gctttgtgagaggt 988
QY 961 caggtgtctctctactactactactactactactactactactactactactactactact 1020
DB 989 caggtgtctctctactactactactactactactactactactactactactactactact 1048
QY 1021 gccggtccagcagcaatccagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1080

Db 1049 gccgcctccagcagcagcattccagatgtgaltctctcttccacaaacagtgcactccc 1108
 Qy 1081 agcatgaccacagattgcacccctccagcaggaacatctctctgacgaatgccaactttg 1140
 Db 1109 agcatgaccacagattgcacccctccagcaggaacatctctctgacgaatgccaactttg 1168
 Qy 1141 tggagctgaccggaagatatagtgtcccaatcccatgtcagtcacatgcacaaactctc 1200
 Db 1169 tggagctgaccggaagatatagtgtcccaatcccatgtcagtcacatgcacaaactctc 1228
 Qy 1201 acctcccaacagccctctcagtgaaacagatgtcaagaacaaactatcaaaagcttccact 1260
 Db 1229 acctcccaacagccctctcagtgaaacagatgtcaagaacaaactatcaaaagcttccact 1288
 Qy 1261 ttaacaaacactgaattctcttccagagctgactcctcc 1298
 Db 1289 ttaacaaacactgaattctcttccagagctgactcctcc 1326

RESULT 3

ID AAF80323 standard; DNA; 1658 BP.
 AC AAF80323;

DT 29-JUN-2001 (first entry)
 XX

DE Nucleotide sequence of G-protein coupled receptor IGS4A short version.
 XX

KM Human: G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrentia;
 KM nervous system disorder; psychiatric disorder; Parkinson's disease;
 KM episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
 KM cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
 KM motility disorder; myocardial infarction; hypertension; dyslipidemia;
 KM gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
 KM inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
 KM gynecological disorder; ss.
 KM

OS Homo sapiens.
 XX

FH Key Location/Qualifiers
 FT CDS 64..1302
 FT /tag= a
 FT /product= "G-protein coupled receptor IGS4A short
 FT version"
 FT

XX MO200125269-A2.
 PN
 XX
 PD 12-APR-2001.
 XX

XX 25-SEP-2000; 2000MO-EP09584.
 PF
 XX

PR 24-SEP-1999; 99EP-0203140.
 PR 24-SEP-1999; 99NL-1013140.
 PR 28-JUL-2000; 2000EP-0202683.
 PR 31-JUL-2000; 2000US-0222047.
 XX

PA (SOLV) SOLVAY PHARM BV.
 XX

PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
 XX

DR WPI: 2001-273568/28.
 XX

PT P-PSDB; AAB67803.
 XX

PT New G-protein coupled receptors and the polynucleotides encoding them,
 PT useful for preventing, ameliorating or correcting nervous system
 PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
 PT or cancers
 XX

PS Claim 1: Page 79-81; 102pp; English.
 XX

CC The present sequence encodes the short version of a human G-protein
 CC coupled receptor designated IGS4A. IGS4 exists in two polymorphic forms,
 CC

CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
 CC for preventing, ameliorating or correcting dysfunctions or diseases.
 CC These diseases include peripheral nervous system, psychiatric and central
 CC nervous system disorders (e.g. schizophrentia, episodic paroxysmal
 CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
 CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
 CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
 CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
 CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
 CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
 CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
 CC effective with regard to disorders of the nervous system, including the
 CC central and peripheral nervous systems, disorders of the gastrointestinal
 CC system, cardiovascular system, skeletal muscle, thyroid, lung or
 CC genitourinary system, or immunological disease. The IGS4 polynucleotides
 CC are useful as diagnostic reagents for detecting under-expression,
 CC overexpression or altered expression of IGS4.
 CC

SQ Sequence 1658 BP; 377 A; 490 C; 346 G; 445 T; 0 other;

Query Match 100.0%; Score 1298; DB 22; Length 1658;
 Best Local Similarity 100.0%; Pred. No. 2,5e-301;
 Matches 1298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 agggagagctcagccttgatttaatgtcaaggatgaaacattcagaatgcttcc 60
 Db 29 agggagagctcagccttgatttaatgtcaaggatgaaacattcagaatgcttcc 88
 Qy 61 ggalctaccagcaaaactagaagatccatccagaacacctgaacagaccaggagt 120
 Db 89 ggalctaccagcaaaactagaagatccatccagaacacctgaacagaccaggagt 148
 Qy 121 atcggcttcctctctggagacctggcgagccactctctcctccggtgtctgtgt 180
 Db 149 atcggcttcctctctggagacctggcgagccactctcctccggtgtctgtgt 208
 Qy 181 atgtgccaatcttctgt 240
 Db 209 atgtgccaatcttctgt 268
 Qy 241 agcacagagctatagaagacgcccaacactactctcttcagcctggcggtctgacc 300
 Db 269 agcacagagctatagaagacgcccaacactactctcttcagcctggcggtctgacc 328
 Qy 301 tccctgtcctctctctggagatgccccttgagagctatgagatgtggcgcaactaccct 360
 Db 329 tccctgtcctctctctggagatgccccttgagagctatgagatgtggcgcaactaccct 388
 Qy 361 tctgttcgggcccgt 420
 Db 389 tctgttcgggcccgt 448
 Qy 421 ccttcacatccatgacatcacacacgttcagcgtgtgagcgtacgttgacatccaccgt 480
 Db 449 ccttcacatccatgacatcacacacgttcagcgtgtgagcgtacgttgacatccaccgt 508
 Qy 481 tccggcgcaaaactgcaagaccccgcgccggcgccctcagaagatccttgagatgtctggg 540
 Db 509 tccggcgcaaaactgcaagaccccgcgccggcgccctcagaagatccttgagatgtctggg 568
 Qy 541 gctctcgtgtctctctcctctgccaacacacacatcatcatgacatcaact 600
 Db 569 gctctcgtgtctctctcctctgccaacacacacatcatcatgacatcaact 628
 Qy 601 tcccaaatggttcctgt 660
 Db 629 tcccaaatggttcctgt 688
 Qy 661 tctacaattcatcatccagatccatctctctcatctacatctctccatgactgtca 720
 Db 689 tctacaattcatcatccagatccatctctctcatctacatctctccatgactgtca 748


```

Oy 721 tcagtcctctactaccctccatcagcactcagactaaagaacaaatctcttgagcag 780
    |||
Db 749 tcagtcctctactaccctccatcagcactcagactaaagaacaaatctcttgagcag 808
Oy 781 atgaaggagatcgaatattcaaaagacccctgcagaaaatcagtaacagatgcttg 840
    |||
Db 809 atgaaggagatcgaatattcaaaagacccctgcagaaaatcagtaacagatgcttg 868
Oy 841 tcttggtcttagtggttgcattatctgttggcccccgttccacatgtgacattcttca 900
    |||
Db 869 tcttggtcttagtggttgcattatctgttggcccccgttccacatgtgacattcttca 928
Oy 901 gcttgtaggagagtgagtgaaatccctctgagctgctgttcaaacctgcacatgtgtgt 960
    |||
Db 929 gcttgtaggagagtgagtgaaatccctctgagctgctgttcaaacctgcacatgtgtgt 988
Oy 961 caggtgtctctctctacactgagctcagctgctcaaccatctatctaaactactgtctc 1020
    |||
Db 989 caggtgtctctctctacactgagctcagctgctcaaccatctatctaaactactgtctc 1048
Oy 1021 gccgccttcagagcagatccccaagatgatactctcttcccaaacagtgacactcc 1080
    |||
Db 1049 gccgccttcagagcagatccccaagatgatactctcttcccaaacagtgacactcc 1108
Oy 1081 agcatgaccacagatgacacactgcccagcagacatcttccctgacagatgacactttg 1140
    |||
Db 1109 agcatgaccacagatgacacactgcccagcagacatcttccctgacagatgacactttg 1168
Oy 1141 tggagctgacagagatagatagctcccaattcccatgctcagatcagtaacaaacttc 1200
    |||
Db 1169 tggagctgacagagatagatagctcccaattcccatgctcagatcagtaacaaacttc 1228
Oy 1201 acctcccaaacagcctctctagtgagacagatgacagaaacaaacttcaaaagcttccact 1260
    |||
Db 1229 acctcccaaacagcctctctagtgagacagatgacagaaacaaacttcaaaagcttccact 1288
Oy 1261 ttaacaaaacctgaattcttccagagctgactctc 1298
    |||
Db 1289 ttaacaaaacctgaattcttccagagctgactctc 1326

RESULT 4
AAF80324
ID AAF80324 standard; DNA; 1658 BP.
XX
XX
AC AAF80324;
XX
XX
DT 29-JUN-2001 (first entry)
XX
XX
DE Nucleotide sequence of a G-protein coupled receptor IGS4B long version.
XX
XX
Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
nervous system disorder; psychiatric disorder; Parkinson's disease;
episodic paroxysmal anxiety; phobia; migraine; bulimia; stroke;
cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
motility disorder; myocardial infarction; hypertension; dyslipidemia;
gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
gynecological disorder; ss.
XX
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 55..1302
FT /tag= a
FT /product= "G-protein coupled receptor IGS4B long version"
XX
XX
PN WO200125269-A2.
XX
XX
PD 12-APR-2001.
XX
XX
PF 25-SEP-2000; 2000WO-EP09584.
XX

```

```

PR 24-SEP-1999: 99EP-0203140.
PR 28-SEP-1999: 99NL-1013140.
PR 28-JUL-2000: 2000EP-0202683.
PR 31-JUL-2000: 2000US-0222047.
XX
XX
PA (SOLV ) SOLVAY PHARM BV.
XX
PI Deleersnijder W, Berger C, Loecken C, Nys G, Venema J:
DR WPI: 2001-273568/28.
DR P-PSDB: AAB67804.
XX
XX
PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers.
XX
XX
PS Claim 1; Page 83-85; 102pp; English.
XX
XX
The present sequence encodes the long version of a human G-protein
coupled receptor designated IGS4B. IGS4 exists in two polymorphic forms,
IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
for preventing, ameliorating or correcting dysfunction or diseases.
These diseases include peripheral nervous system, psychiatric and central
nervous system disorders (e.g. schizophrenia, episodic paroxysmal
anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
gastrointestinal disorders (e.g. inflammatory bowel disease or motility
disorders), osteoporosis, inflammations, infections (e.g. bacterial,
fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
effective with regard to disorders of the nervous system, including the
central and peripheral nervous systems, disorders of the gastrointestinal
system, cardiovascular system, skeletal muscle, thyroid, lung or
gastrointestinal system, or immunological disease. The IGS4 polynucleotides
are useful as diagnostic reagents for detecting under-expression,
overexpression or altered expression of IGS4.
XX
XX
Sequence 1658 BP; 378 A; 489 C; 346 G; 445 T; 0 other:
SO

```

Query Match 99.5%; Score 1291.6; DB 22; Length 1658;
 Best Local Similarity 99.7%; Pred. No. 8.4e-300;
 Matches 1294; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

Oy 1 agggagagctcagccttgatttaatgctcagagatgagaaaactcagaatgctcct 60
    |||
Db 29 agggagagctcagccttgatttaatgctcagagatgagaaaactcagaatgctcct 88
Oy 61 ggaatcaccagcagaactagaatgccatccagaagaaccccgaaacagcaggagagt 120
    |||
Db 89 ggaatcaccagcagaactagaatgccatccagaagaaccccgaaacagcaggagagt 148
Oy 121 atctggcttcctctcgagaccctcgagcagcaactctcccccgtctgtgtgtg 180
    |||
Db 149 atctggcttcctctcgagaccctcgagcagcaactctcccccgtctgtgtgtg 208
Oy 181 atgtgccaatcttgctgtgtgtgtgcatgtgcaatgctcctgtgtgtgtgtgtc 240
    |||
Db 209 atgtgccaatcttgctgtgtgtgtgcatgtgcaatgctcctgtgtgtgtgtgtc 268
Oy 241 agacacaggtatagaagagcccaacatacactctcagcctggcggtctctgagc 300
    |||
Db 269 agacacaggtatagaagagcccaacatacactctcagcctggcggtctctgagc 328
Oy 301 tcttggtcctgctccttgagatgcccttgagagttcatatgagatgttggcgaactaccct 360
    |||
Db 329 tcttggtcctgctccttgagatgcccttgagagttcatatgagatgttggcgaactaccct 388
Oy 361 tcttggtcggcccgctggagctgctactcaagacggccctcttgagaccgtgtgtctcg 420
    |||
Db 389 tcttggtcggcccgctggagctgctactcaagacggccctcttgagaccgtgtgtctcg 448

```

Qy	421	gcctcatctccagatcatcaacacgctcaagcgttgagagcgtctagcttgagcatctctaacccgct	480
Dp	449	ccctccacacctccatagatcatcaacacgctccagctgagagcgtctagcgttgagcatctctaacccgct	508
Qy	481	tcagcgcgcacaaactctgacagacgacccgcgcgcgcgcctctcagatctccctgagcatctgctg	540
Dp	509	tcagcgcgcacaaactctgacagagacccgcgcgcgcgcctctcagatctccctgagcatctgctg	568
Qy	541	gctctctccgctgctctctctccctctgcccacacacgacatccatgagcatcaagtctccactact	600
Dp	569	gctctctccgctgctctctctccctctgcccacacacgacatccatgagcatcaagtctccactact	628
Qy	601	tcgcccaatctgagctccctctgctccagagcttcagacccacccctgctacgtcatcaagcccatctg	660
Dp	629	tcgcccaatctgagctccctctgctccagagcttcagacccacccctgctacgtcatcaagcccatctg	688
Qy	661	tcctacaattctcatccacaggtcaacctctctctatctcaactctctcccatctgactgtca	720
Dp	689	tcctacaattctcatctcatccaggtcaacctctctctatctcaactctctcccatctgactgtca	748
Qy	721	tcagatgctccctctatctacatctgacactctcaagatctaaagaaagaaatctctctggagcag	780
Dp	749	tcagatgctccctctatctacatctgacactctcaagatctaaagaaagaaatctctctggagcag	808
Qy	781	atggaagaggaatctgacaaatatctcaaaagacccctctgacgaaatactagctcaaaagatctgct	840
Dp	809	atggaagaggaatctgacaaatatctcaaaagacccctctgacgaaatactagctcaaaagatctgct	868
Qy	841	tccttgagctcttagctgtctgctcatctgtctgggcccgcgtctcaacattctgacgacactctctca	900
Dp	869	tccttgagctcttagctgtctgctcatctgtctgggcccgcgtctcaacattctgacgacactctctca	928
Qy	901	gctcttgctgagaggaatctgagatgtaactccctgagctgctgtgtctcaaacctctgacatctgct	960
Dp	929	gctcttgctgagaggaatctgagatgtaactccctgagctgctgtgtgtctcaaacctctgacatctgct	988
Qy	961	caggtgctctctctctacactgagctctcagctgtcaaaccccatctatctataacactgctctc	1020
Dp	989	caggtgctctctctctctacactgagctctcagctgtgtcaaaccccatctatctataacactgctctc	1048
Qy	1021	gcgcgctctccagacgacatctccagaaatgtaactctctctctccacaaacagctgacactccc	1080
Dp	1049	gcgcgctctccagacgacatctccagaaatgtaactctctctctctccacaaacagctgacactccc	1108
Qy	1081	agcatctgaccccaagctctgacacactgcccacgcggaatactctctgacagaaatgacactctg	1140
Dp	1109	agcatctgaccccaagctctgacacactgcccacgcggaatactctctgacagaaatgacactctg	1168
Qy	1141	ctgagagctcgaaacggaagataatggtctcccaattctcccatgtaatgcatctccatgacaaactctc	1200
Dp	1169	ctgagagctcgaaacggaagataatggtctcccaattctcccatgtaatgcatctccatgacaaactctc	1228
Qy	1201	acctctccaaacagccctctctctagctgtaaacagatatctcaagaaacaaactctcaagctctccact	1260
Dp	1229	acctctccaaacagccctctctctagctgtaaacagatatctcaagaaacaaactctcaagctctccact	1288
Qy	1261	cttaacaaaaacttgatattctctctcagagctgagctctctc	1298
Dp	1289	cttaacaaaaacttgatattctctctcagagctgagctctctc	1326

RESULT	5
AAFB0325	
ID	AAFB0325 standard; DNA; 1658 BP.
XX	
AC	AAFB0325;
XX	
DT	29-JUN-2001 (first entry)
XX	
DE	Nucleotide sequence of G-protein coupled receptor IGS4B short version.
XX	
XX	Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;

	XX	nervous system disorder; psychiatric disorder; Parkinson's disease;
KW	epistodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;	
KM	cardiovascular disease; heart failure; angina pectoris; obesity; emesis;	
KV	motility disorder; myocardial infarction; hypertension; dyslipidemia;	
KW	gastrointestinal disorder; inflammatory bowel disease; osteoporosis;	
KW	inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;	
KW	gynecological disorder; ss.	
OS	Homo sapiens.	
XX		
FH	Key Location/Qualifiers	
FT	CDS 64..1302	
FT	/tag= a	
FT	/product= "G-protein coupled receptor IGS4B short version"	
FT		
XX		
PX	WO200125269-A2.	
PD		
PD	12-Apr-2001.	
XX		
PF	25-SEP-2000; 200OWO-EPO9584.	
PR	24-SEP-1999; 99EP-0203140.	
PR	24-SEP-1999; 99NL-1013140.	
PR	28-JUL-2000; 200OE-0220683.	
PR	31-JUL-2000; 200OU-02202047.	
PA	(SOLV) SOLVAY PHARM BV.	
XX		
PI	Deleersnijder W, Berger C, Loeken C, Nys G, Venema J:	
DR	WPI: 2001-273568/28.	
XX		
DRI	P-PSDB: AAB67805.	
PT	New G-protein coupled receptors and the polynucleotides encoding them,	
PT	useful for preventing, ameliorating or correcting nervous system	
PT	disorders, cardiovascular diseases, dyslipidemias, inflammations, pain	
PT	or cancers -	
XX		
PS	Claim 1; Page 87-89; 102pp; English.	
CC	The present sequence encodes the short version of a human G-protein	
CC	coupled receptor designated IGS4B. IGS4 exists in two polymorphic forms,	
CC	IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful	
CC	for preventing, ameliorating or correcting dysfunction disorders.	
CC	These diseases include peripheral nervous system, psychiatric and central	
CC	nervous system disorders (e.g. schizophrenia), episodic paroxysmal	
CC	anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or	
CC	stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,	
CC	myocardial infarction or hypertension), dyslipidemias, obesity, emesis,	
CC	gastrointestinal disorders (e.g. inflammatory bowel disease or motility	
CC	disorders), osteoporosis, inflammations, infections (e.g. bacterial,	
CC	fungal, protozoan or viral), pain, cancers, immune disorders, allergies,	
CC	sepsis or gynecological disorders. Agonists or antagonists of IGS4 are	
CC	effective with regard to disorders of the nervous system, including the	
CC	central and peripheral nervous systems, disorders of the gastrointestinal	
CC	system, cardiovascular system, skeletal muscle, thyroid, lung or	
CC	genitourinary system, or immunological disease. The IGS4 polynucleotides	
CC	are useful as diagnostic reagents for detecting under-expression,	
CC	overexpression or altered expression of IGS4.	
SQ		
	Sequence 1658 BP; 378 A; 489 C; 346 G; 445 T; 0 other;	
Query Match	99.5%; Score 1291.6; DB 22; Length 1658;	
Best Local Similarity	99.7%; Pred. No. 8.4e-300;	
Matches 1294; Conservative	0; Mismatches 4; Indels 0; Gaps 0;	
OY	I agggagaggctcagcgcttgatattcaibgttcagggatgaanaacttcagaatgccttcc 60 	
Db	29 agggagaggctcagcgcttgatattcaibgttcagggatgaanaacttcagaatgccttcc 88	
OY	61 ggatcacacagaagaactagaaagataccaattcccagaacaaccttgacacagcgagagt 120	

```

Db      89 ggaattacacagacagaacatagatgccattccagaacacctgacacagcaggagt 148
Oy      121 atctggcctctctctgcggaacctgcgcgagcaccattctctcccggtctggtgt 180
Db      149 atctggcctctctctgcggaacctgcgcgagcaccattctctcccggtctggtgt 208
Oy      181 atgtgccaaattctgtggtggtggtcatgtgcaatgtctctgtgtggtccgtggaatctgc 240
Db      209 atgtgccaattctgtgtggtggtggtcatgtgcaatgtctctgtgtggtccgtggaatctgc 268
Oy      241 agcaccagagctatgaagacgcccacaaactactacctcttcagcctgycgtctctgacc 300
Db      269 agcaccagagctatgaagacgcccacaaactactacctcttcagcctgycgtctctgacc 328
Oy      301 tctgtgtctctgtctctctggaatgccccctggaagtctctatgaatgtggtgcaactacctt 360
Db      329 tctgtgtctctgtctctctggaatgccccctggaagtctctatgaatgtggtgcaactacctt 388
Oy      361 tctgtgtctctgtctctctggaatgccccctggaagtctctatgaatgtggtgcaactacctt 420
Db      389 tctgtgtctctgtctctctggaatgccccctggaagtctctatgaatgtggtgcaactacctt 448
Oy      421 cctccatctctcagcatcaccaacgctcaagcgtggaagcgtctggtgcaactacctt 480
Db      449 cctccatctctcagcatcaccaacgctcaagcgtggaagcgtctggtgcaactacctt 508
Oy      481 tccggtcccaaatctgacagagcaccggtcgcggtccctcagatctcccggtcgtgtgtgt 540
Db      509 tccggtcccaaatctgacagagcaccggtcgcggtccctcagatctcccggtcgtgtgtgt 568
Oy      541 gcttctcgtgtctctctctcgtcccaacacacagcatctcatgcatcaggtctccactact 600
Db      569 gcttctcgtgtctctctctcgtcccaacacacagcatctcatgcatcaggtctccactact 628
Oy      601 tccccaatggtgtccctgtgtcccaaggtctcggtcaccgtgtacgtgtccatcaagcccatgtgga 660
Db      629 tccccaatggtgtccctgtgtcccaaggtctcggtcaccgtgtacgtgtccatcaagcccatgtgga 688
Oy      661 tctaaacttctctctccaggtgtccatctctctctctctctctctccatgactgtgtca 720
Db      689 tctaaacttctctctccaggtgtccatctctctctctctctctctccatgactgtgtca 748
Oy      721 tcaagtgtctctactacacatcgtgcatcagactaagaagaacaaatctctgtgaagcag 780
Db      749 tcaagtgtctctactacacatcgtgcatcagactaagaagaacaaatctctgtgaagcag 808
Oy      781 atgaagggaatgtcaaatatccaagaagccctgacagaataatcaatcaacaagaatgtgtttg 840
Db      809 atgaagggaatgtcaaatatccaagaagccctgacagaataatcaatcaacaagaatgtgtttg 868
Oy      841 tctgtgtctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 900
Db      869 tctgtgtctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 928
Oy      901 gcttctggaagaggtggaatgtgaatccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
Db      929 gcttctggaagaggtggaatgtgaatccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 988
Oy      961 caagtgtctctctctactacatcagactcagactgttcaaccccaatctctatcaactactgtctc 1020
Db      989 caagtgtctctctctactacatcagactcagactgttcaaccccaatctctatcaactactgtctc 1048
Oy      1021 gctgtgtctcaggtcagcatctccagaatgtgactctctctctctccaaaacaggtgtgacctcc 1080
Db      1049 gctgtgtctcaggtcagcatctccagaatgtgactctctctctctccaaaacaggtgtgacctcc 1108
Oy      1081 agcatgaccacagatgtgacactgtccagcaggaatactctctctgacagagtccactttg 1140
Db      1109 agcatgaccacagatgtgacactgtccagcaggaatactctctctgacagagtccactttg 1168
Oy      1141 tggagctgacccgaagatataaggttccccaatcccaatgtcagatcatalcagacaaactctc 1200

```

```

Db      1169 tggagctgacccgaagatataaggtccccaattcctatgtctcagtcacgtgtgacaaactctc 1228
Oy      1201 acctcccaacagccctctctatgtgacagatgttcaagaacaaactataaagcttccact 1260
Db      1229 acctcccaacagccctctctatgtgacagatgttcaagaacaaactataaagcttccact 1288
Oy      1261 ttaacaaacctgaattctcttcagagctgactctctc 1298
Db      1289 ttaacaaacctgaattctcttcagagctgactctctc 1326

RESULT 6
AA233297
ID AA233297 standard; cDNA; 1360 BP.
XX
AC AA233297;
XX
DT 21-FEB-2000 (first entry)
XX
DE Human neurotensin-like receptor encoding cDNA.
XX
KW Human; neurotensin-like receptor; NLR: G-protein coupled receptor;
KW central nervous system; anesthesia; analgesia; ds.
XX
OS Homo sapiens.
XX
PN W09955732-A1.
XX
PD 04-NOV-1999.
XX
PF 15-APR-1999; 99WO-SE00598.
XX
PR 24-APR-1998; 98SE-0001455.
XX
PA (ASTR-) ASTRA PHARMA INC.
PA (ASTR) ASTRA AB.
XX
PI Ahmad S, Cao J, O'Donnell D, Walker P;
XX
DR WPI. 2000-052803/04.
DR P-PSDB; AAY52992.
XX
PT Novel neurotensin-like receptor, useful for identifying agents for
PT producing anesthesia or analgesia
XX
PS Claim 11; Fig 1; 46pp; English.
XX
CC The present sequence encodes a human G-protein coupled receptor
CC neurotensin-like receptor (NLR). The NLR polynucleotide and protein
CC can be used to isolate compounds that bind, (ant)agonise or alter
CC the activity or expression of the NLR. The NLR is a G-protein coupled
CC receptor which is expressed in the central nervous system and shares
CC homology with human neurotensin receptor. The receptors can be used in
CC assays to identify agents for producing anesthesia and analgesia.
XX
SO Sequence 1360 BP; 293 A; 423 C; 294 G; 350 T; 0 other;

Query Match 99.1%; Score 1286.8; DB 21; Length 1360;
Best Local Similarity 99.5%; Pred. No. 1.le-298;
Matches 1291; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy      1 agggaggtcaggtcaggtcgtgatttaagtctcagggatggaataaacttcagaatgtctctc 60
Db      39 agggaggtcaggtcaggtcgtgatttaagtctcagggatggaataaacttcagaatgtctctc 98
Oy      61 ggatctacagcagaacactagaagatccattccagaanaacacttgacagcaccggaggagt 120
Db      99 ggaattacacagacagaacactagaagatccattccagaanaacacttgacagcaccggaggagt 158
Oy      121 atctggtctctctctggtgacactcggtcagcgaactctctctcccggtgtgtgtgtgt 180
Db      159 atctggtctctctctggtgacactcggtcagcgaactctctctcccggtgtgtgtgtgt 218

```

[illegible]

Oy	1261	ttacaacaaactgaattcttccagagctgacctcctc	1298
Dd	1299	ttacaacaaactgaattcttccagagctgacctcctc	1336
RESULT# 7			
AD01123	ID	AD01123 standard; cDNA; 1248 BP.	
XX AC		AD01123;	
XX DT	02-NOV-2000	(first entry)	
XX DE		Human orphan G protein-coupled receptor hrUP6 cDNA.	
XX KW		Human; orphan G protein-coupled receptor; GPCR; hrUP6; drug screening;	
XX XX		transmembrane receptor; signal cascade; ss.	
OS		Homo sapiens.	
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..1248	
FT		/*tag= a	
FT		/product= "hrUP6"	
FT		/note= "human orphan G protein-coupled receptor"	
XX PN		WO200031258-A2.	
XX PD	02-JUN-2000.		
XX PF	13-OCT-1999;	99WO-US23687.	
XX PR			
PR	20-NOV-1998;	98US-0109213.	
PR	16-FEB-1999;	99US-0120416.	
PR	26-FEB-1999;	99US-0121852.	
PR	12-MAR-1999;	99US-0123946.	
PR	12-MAR-1999;	99US-0123949.	
PR	28-MAY-1999;	99US-0136436.	
PR	28-MAY-1999;	99US-0136437.	
PR	28-MAY-1999;	99US-0136439.	
PR	28-MAY-1999;	99US-0136567.	
PR	28-MAY-1999;	99US-0137127.	
PR	28-MAY-1999;	99US-0137131.	
PR	29-JUN-1999;	99US-0141448.	
PR	29-SEP-1999;	99US-0156555.	
PR	29-SEP-1999;	99US-0156653.	
PR	29-SEP-1999;	99US-0156654.	
PR	29-SEP-1999;	99US-0156657.	
PR	01-OCT-1999;	99US-0157280.	
PR	01-OCT-1999;	99US-0157281.	
PR	01-OCT-1999;	99US-0157282.	
PR	01-OCT-1999;	99US-0157293.	
PR	01-OCT-1999;	99US-0157294.	
PR	12-OCT-1999;	99US-0416760.	
PR	12-OCT-1999;	99US-0417044.	
XX PA		(AREN-) ARENA PHARM INC.	
XX PI			
PI	Chen R, Dang HT, Liaw CW, Lin I,		
DR	WPI: 2000-400068/34.		
DR	P-PsDB; MAY71296.		
XX PT			
PT	Novel human orphan G protein-coupled receptors and the encoding cDNAs		
CC	for use in the identification of G protein-coupled receptor agonists -		
XX PS	Claim 21; Page 56-57; 102pp; English.		
XX CC			
CC	The present sequence is a cDNA encoding hrUP6, an endogenous human		
CC	orphan G protein-coupled receptor (GPCR). The full length hrUP6 cDNA		
CC	was cloned by RT-PCR using human thymus cDNA as template.		
CC	The orphan GPCR of the invention, like all GPCRs has seven transmembrane		
CC	alpha helices with an extracellular N-terminus and an intracellular		

CC C-terminus. However, no endogenous ligands has yet been identified for
CC the proteins of the invention. The orphan GPCRs may be used in the
CC identification of their endogenous ligands, and to screen potential GPCR
CC agonists and antagonists for use as pharmaceutical agents. The proteins
CC may also be used in the study of GPCR-mediated signalling cascades, and
CC to elucidate their precise role in normal and diseased human conditions.
CC Nucleic acid encoding human orphan GPCRs may be used for tissue
CC localisation expression analysis to provide information about their
CC function in healthy and pathological states.
CC
XX

SQ Sequence 1248 BP; 269 A; 393 C; 269 G; 317 T; 0 other:

Query Match 96.1%; Score 1248; DB 21; Length 1248;
Best Local Similarity 100.0%; Pred. No. 2,1e-289;
Matches 1248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 27 atgtcagagatggaanaacttcagaatgctctctgatactacagagaactagaat 86
DB 1 atgtcagagatggaanaacttcagaatgctctctgatactacagagaactagaat 60
OY 87 ccatccagaagaacacctggaacagacagagagatctgctctctcggaacctcg 146
DB 61 ccatccagaagaacacctggaacagacagagagatctgctctctcggaacctcg 120
OY 147 cgcagcgaactctctctcccgctgctggtgatactgacaaatttttggtgggggtc 206
DB 121 cgcagcgaactctctctcccgctgctggtgatactgacaaatttttggtgggggtc 180
OY 207 attggaacatgctctgctgctgctgatacttcgacagacagagcctaagaagccacc 266
DB 181 attggaacatgctctgctgctgctgatacttcgacagacagagcctaagaagccacc 240
OY 267 aactataactcttcaagcctgctgctgctgatacttcgacagacagagcctaagaagcc 326
DB 241 aactataactcttcaagcctgctgctgctgatacttcgacagacagagcctaagaagcc 300
OY 327 cggagagctatagatgctgctgctgctgatacttcgacagacagagcctaagaagcc 386
DB 301 cggagagctatagatgctgctgctgctgatacttcgacagacagagcctaagaagcc 360
OY 387 ttcaagaagcggccctcttgaagacgctgctgctgatacttcgacagacagagcctaaga 446
DB 361 ttcaagaagcggccctcttgaagacgctgctgctgatacttcgacagacagagcctaaga 420
OY 447 agcgtgagagcgtatagctgctgctgatacttcgacagacagagcctaagaagcc 506
DB 421 agcgtgagagcgtatagctgctgctgatacttcgacagacagagcctaagaagcc 480
OY 507 cgcgagccctcagagatctctgctgctgctgatacttcgacagacagagcctaagaagcc 566
DB 481 cgcgagccctcagagatctctgctgctgctgatacttcgacagacagagcctaagaagcc 540
OY 567 aacacacagatccatggaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 626
DB 541 aacacacagatccatggaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 600
OY 627 tcggcgaacctgtgaagctatcgaagccatgctgatacttcgaatcgaatcgaatcgaat 686
DB 601 tcggcgaacctgtgaagctatcgaagccatgctgatacttcgaatcgaatcgaatcgaat 660
OY 687 tccctcctatctacatctctcccaatgatactgatactgatactgatactgatactgata 746
DB 661 tccctcctatctacatctctcccaatgatactgatactgatactgatactgatactgata 720
OY 747 ctcaagaactaagaagaacaaatctcttgaagcagatgaagaaggaatgaataatcaaaaga 806
DB 721 ctcaagaactaagaagaacaaatctcttgaagcagatgaagaaggaatgaataatcaaaaga 780
OY 807 cccctgcagaaaaatcagtaacaagaatgctgctgctgctgctgctgctgctgctgctgct 866
DB 781 cccctgcagaaaaatcagtaacaagaatgctgctgctgctgctgctgctgctgctgctgct 840
```

```
OY 867 tgggcccttcacatctgacagactctctcagctctgtgagagatggaatgaccc 926
DB 841 tgggcccttcacatctgacagactctctcagctctgtgagagatggaatgaccc 900
OY 927 ctgctgctgtgttcaacctctgacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 986
DB 901 ctgctgctgtgttcaacctctgacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
OY 987 gctgttaaccccatctatataacctactgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1046
DB 961 gctgttaaccccatctatataacctactgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1020
OY 1047 gtagctctctcttcccaaacagtgacactctccagcagatgacacagatgacacagatgac 1106
DB 1021 gtagctctctcttcccaaacagtgacactctccagcagatgacacagatgacacagatgac 1080
OY 1107 cagcggaacatctctctgacagaaatgcaactctgtggaagctgacagcgaagataagatccc 1166
DB 1081 cagcggaacatctctctgacagaaatgcaactctgtggaagctgacagcgaagataagatccc 1140
OY 1167 caattccatgtcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcat 1226
DB 1141 caattccatgtcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcat 1200
OY 1227 cagatgtcaagaacaacatcatcaagcttccacttcaacaaacctga 1274
DB 1201 cagatgtcaagaacaacatcatcaagcttccacttcaacaaacctga 1248

RESULT 8
ID AAA46022
AAA46022 standard; cDNA; 1248 BP.
XX
AC AAA46022:
XX
DT 22-AUG-2000 (first entry)
XX
DE Human G protein coupled receptor hRup6 encoding cDNA SEQ ID NO:11.
KW Human; G protein coupled receptor; GPCR; transmembrane receptor;
KW Identification; agonist; screening; therapeutic; pharmaceutical;
XX mutant; ss.
OS Homo sapiens.
XX
PN WO200022131-A2.
XX
PD 20-APR-2000.
XX
PF 13-OCT-1999; 99WO-US24065.
XX
PR 13-OCT-1998; 98US-0170496.
XX
PR 12-NOV-1998; 98US-0108029.
XX
PR 20-NOV-1998; 98US-0109213.
XX
PR 27-NOV-1998; 98US-0110060.
XX
PR 16-FEB-1999; 99US-0120416.
XX
PR 26-FEB-1999; 99US-0121852.
XX
PR 12-MAR-1999; 99US-0123944.
XX
PR 12-MAR-1999; 99US-0123945.
XX
PR 12-MAR-1999; 99US-0123946.
XX
PR 12-MAR-1999; 99US-0123948.
XX
PR 12-MAR-1999; 99US-0123949.
XX
PR 12-MAR-1999; 99US-0123951.
XX
PR 12-MAR-1999; 99US-0123951.
XX
PR 28-MAY-1999; 99US-0136436.
XX
PR 28-MAY-1999; 99US-0136437.
XX
PR 28-MAY-1999; 99US-0136439.
XX
PR 28-MAY-1999; 99US-0137127.
XX
PR 28-MAY-1999; 99US-0137131.
XX
PR 28-MAY-1999; 99US-0137131.
XX
PR 28-MAY-1999; 99US-0137567.
XX
PR 30-JUN-1999; 99US-0141448.
XX
PR 27-AUG-1999; 99US-0151114.
XX
PR 03-SEP-1999; 99US-0152524.
XX
PR 29-SEP-1999; 99US-0156633.
```

PR 29-SEP-1999; 9905-0156555.
 PR 29-SEP-1999; 9905-0156634.
 XX
 PA (AREN-) ARENA PHARM INC.
 XX

PI Behan DP, Lehmann-Brulisma K, Chalmers DT, Chen R, Dang HT;
 PI Gore M, Liaw CW, Lin I, Lowitz K, White C;
 XX
 DR WPI: 2000-317986/27.
 XX P-PSDB: AAB02830.

PT Non-endogenous, human G protein-coupled receptors for screening
 PT receptor, inverse or partial agonists useful as therapeutic agents
 PS
 XX
 PS Example 1; Page 85-86; 187pp: English.

CC The present invention describes transmembrane receptors, preferably
 CC human G protein coupled receptors (GPCR), for which the endogenous
 CC ligand is unknown (orphan GPCR receptors). More specifically the present
 CC invention relates to non-endogenous, constitutively activated versions
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for
 CC the direct identification of candidate compounds as receptors agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents.
 CC AAA6017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 CC the exemplification of the present invention.
 XX
 XX Sequence 1248 BP; 269 A; 393 C; 269 G; 317 T; 0 other;

Query Match 96.1%; Score 1248; DB 21; Length 1248;
 Best Local Similarity 100.0%; Pred. No. 2,1e-289;
 Matches 1248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 atgtcagggatggaagaaacttcagaatgctccttgatctacacaggaagaaatagaat 85
 Db 1 atgtcagggatggaagaaacttcagaatgctccttgatctacacaggaagaaatagaat 60
 QY 87 ccaatccagaagaacacctgaaacagcagcagagagatctgagcctctctcgtgagacctcg 146
 Db 61 ccaatccagaagaacacctgaaacagcagcagagagatctgagcctctctcgtgagacctcg 120
 QY 147 cgaagcgaactctctctcccggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 206
 Db 121 cgaagcgaactctctctcccggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 180
 QY 207 attggaatgtcctgt 266
 Db 181 attggaatgtcctgt 240
 QY 267 aacttaacctcttcagcctgagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 326
 Db 241 aacttaacctcttcagcctgagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 300
 QY 327 ctggagtgctatagatgt 386
 Db 301 ctggagtgctatagatgt 360
 QY 387 tcaaaagcggccctctcttgagacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 446
 Db 361 tcaaaagcggccctctcttgagacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 420
 QY 447 agcgtgagagcgtacgt 506
 Db 421 agcgtgagagcgtacgt 480
 QY 507 cgcggggccctcaggaatcctcggcagatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 566
 Db 481 cgcggggccctcaggaatcctcggcagatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 540
 QY 567 aacacacagcatcagcatcagcatcagcatcagcatcagcatcagcatcagcatcagcatcag 626
 Db 541 aacacacagcatcagcatcagcatcagcatcagcatcagcatcagcatcagcatcagcatcag 600

QY 627 tcggcacccgtatagatcatcaagcccatgttgatctacatctcatcatccaggtcacc 686
 Db 601 tcggcacccgtatagatcatcaagcccatgttgatctacatctcatcatccaggtcacc 660
 QY 687 tcccttctatcttaacctctccccaatgactgtcatcaggtgtcctctactactactatgca 746
 Db 661 tcccttctatcttaacctctccccaatgactgtcatcaggtgtcctctactactactatgca 720
 QY 747 ctcaagactaaagaagaagaagaatctcttgagcagatgaaaggaatgaaatattcaaga 806
 Db 721 ctcaagactaaagaagaagaagaatctcttgagcagatgaaaggaatgaaatattcaaga 780
 QY 807 cccctgcagaaatcagtcacaaagatgtctgttctgttctgttctgttctgttctgttctgt 866
 Db 781 cccctgcagaaatcagtcacaaagatgtctgttctgttctgttctgttctgttctgttctgt 840
 QY 867 tgggcccgttccaatctgacacgactctcttccaagcttctgtgagagatgagatgaatcc 926
 Db 841 tgggcccgttccaatctgacacgactctcttccaagcttctgtgagagatgagatgaatcc 900
 QY 927 ctgagctgtgttccaacctcgcagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 986
 Db 901 ctgagctgtgttccaacctcgcagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
 QY 987 gctgtcaaccatctatctataactactgtctcgcgcgttccagagcagatccagaat 1046
 Db 961 gctgtcaaccatctatctataactactgtctcgcgcgttccagagcagatccagaat 1020
 QY 1047 gtatctctctcttccaagaacagtgagactcccaagcagatgacacaggttgacactgtcc 1106
 Db 1021 gtatctctctcttccaagaacagtgagactcccaagcagatgacacaggttgacactgtcc 1080
 QY 1107 cagcggagaactcttccagagaatgacacttctgtgagcttgacggagaatagatgtccc 1166
 Db 1081 cagcggagaactcttccagagaatgacacttctgtgagcttgacggagaatagatgtccc 1140
 QY 1167 caattccatgtcagtcacatccatccatccatccatccatccatccatccatccatccat 1226
 Db 1141 caattccatgtcagtcacatccatccatccatccatccatccatccatccatccatccat 1200
 QY 1227 caagatgtcaagaacaaactatcaaaagcttccacttcaaaaacacttga 1274
 Db 1201 caagatgtcaagaacaaactatcaaaagcttccacttcaaaaacacttga 1248

RESULT 9
 AAH49526
 ID AAH49526 standard; DNA; 1248 BP.
 XX
 AC AAH49526;
 XX
 DT 21-SEP-2001 (first entry)
 XX
 DE Human GTP-binding protein-coupled receptor GPRV39 coding sequence.
 XX
 KW GTP-binding protein-coupled receptor; neuroprotective; immunomodulatory;
 KW muscular; urinary; circulatory; anorectic; gene therapy; human;
 KW guanosine triphosphate; G-protein; ds.
 OS Homo sapiens.
 XX
 PN WO200148189-A1.
 XX
 PD 05-JUL-2001.
 XX
 PF 28-DEC-2000; 2000MO-JP09409.
 XX
 PR 28-DEC-1999; 99JP-0375152.
 PR 31-MAR-2000; 2000JP-0101339.
 PR 23-MAY-2000; 2000JP-0155978.
 XX
 PA (HELI-) HELIX RES INST.
 XX

PI Okubo S:
 XX WPI: 2001-488917/53.
 DR P-PSDB: AAG63353.
 XX
 PT Identifying predicted or actual structures of two or more members of a
 PT chemical or physical library by mass spectrometry compiling
 PT correlating molecular mass measurements of two or more members with a
 PT shared chemical history -
 XX
 PS Claim 9; Page 79-80; 95pp: Japanese.
 XX
 CC The present sequence encodes a human TGR-1 protein. The specification
 CC describes a method of screening a compound, which is capable of binding
 CC properties of neuromedin U to TGR-1. The method is useful for screening
 CC preventatives and remedies for hypertension, stress diseases, etc..
 CC TGR-1 antagonists are also useful for treating the same diseases.
 CC
 XX
 SQ Sequence 1245 BP; 268 A; 393 C; 268 G; 316 T; 0 other;

Query Match 95.9%; Score 1245; DB 22; Length 1245;
 Best Local Similarity 100.0%; Pred. No. 1,1e-288;
 Matches 1245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 atgtcaggagtggaagaaacttcagatgcttcttgagatctaccagcagaaactagaagat 85
 Db 1 atgtcaggagtggaagaaacttcagatgcttcttgagatctaccagcagaaactagaagat 60
 QY 87 ccattccagaagaacacttgaacacagcagcagagagataltctgaccttctctctgagacctcg 146
 Db 61 ccattccagaagaacacttgaacacagcagcagagagataltctgaccttctctctgagacctcg 120
 QY 147 cgcagcgaacttctctctcccggtgtctgtgtatgtatgtccaattttctgtgtggtgtc 206
 Db 121 cgcagcgaacttctctctcccggtgtctgtgtatgtatgtccaatlttctgtgtggtgtc 180
 QY 207 attcgcaatgtctcctgtgtctcctgtgtatcttcagacacagcagatgaagacgccacc 266
 Db 181 attcgcaatgtctcctgtgtctcctgtgtatcttcagacacagcagatgaagacgccacc 240
 QY 267 aactaactaactcttcagccttgcgtgtcttcgaacctctgtctctgtccttgaatgccc 326
 Db 241 aactaactaactcttcagccttgcgtgtcttcgaacctctgtcctgtccttgaatgccc 300
 QY 327 ctggagagtcataagatgtgtgcgaactaactcttcttctgttgcggcccggtggtctgac 386
 Db 301 ctggagagtcataagatgtgtgcgaactaactcttcttctgttgcggcccggtggtctgac 360
 QY 387 ttcaagaagcgcctcttctgagacgtgtgtctgcctctcaaccctcagacacacacgctc 446
 Db 361 ttcaagaagcgcctcttctgagacgtgtgtctgcctctcaaccctcagacacacacgctc 420
 QY 447 agcgttgagcgtctatgtgcacactcaaccgcgttccgcgcgaactgtcagagaccccg 506
 Db 421 agcgttgagcgtctatgtgcacactcaaccgcgttccgcgcgaactgtcagagaccccg 480
 QY 507 cgcgcggagccttcagatccctcgcgcgcatgtctgtggtcttcgcgtgtccttccctgccc 566
 Db 481 cgcgcggagccttcagatccctcgcgcgcatgtctgtggtcttcgcgtgtccttccctgccc 540
 QY 567 aacacagcgaatcgaatgcaacaagtctcacaacttcccaatgggtccctggtccagatc 626
 Db 541 aacacagcgaatcgaatgcaacaagtctcacaacttcccaatgggtccctggtccagatc 600
 QY 627 tcggcgaactgtatagtcataagcccaatgttgatctacaatltcatcatccaggtaccc 686
 Db 601 tcggcgaactgtatagtcataagcccaatgttgatctacaatltcatcatccaggtaccc 660
 QY 687 tctctctatcttaactctccctcccaatgactgatacagtgcttctctctactctccatgga 746
 Db 661 tctctctatcttaactctccctcccaatgactgatacagtgcttctctctactctccatgga 720

QY 747 ctcaagaataaagaagacaaatctcttgagcagatgaaggaatgcaaatatccaaga 806
 Db 721 ctcaagaataaagaagacaaatctcttgagcagatgaaggaatgcaaatatccaaga 780
 QY 807 cccctcagaagaatcagtcacaagaatgtgtgtcttggtcttaatgtttgtatctgt 866
 Db 781 cccctcagaagaatcagtcacaagaatgtgtgtcttggtcttaatgtttgtatctgt 840
 QY 867 tgggccccgttcacacattgacagcactcttcttaagcttttgagagatggaatgaccc 926
 Db 841 tgggccccgttcacacattgacagcactcttcttaagcttttgagagatggaatgaccc 900
 QY 927 ctggctgtgtgttcaacctgcgtccatgtgtgtgtcagtggtcttcttcaactgaacctca 986
 Db 901 ctggctgtgtgttcaacctgcgtccatgtgtgtgtcagtggtcttcttcaactgaacctca 960
 QY 987 gctgtcaaacccatctatctaactactgtctcgcgcgttccagcagcatccagaat 1046
 Db 961 gctgtcaaacccatctatctaactactgtctcgcgcgttccagcagcatccagaat 1020
 QY 1047 gtatctcttcttccacaagaatgtgcaactctccagcagatgacccagatgtgccactgccc 1106
 Db 1021 gtatctcttcttccacaagaatgtgcaactctccagcagatgacccagatgtgccactgccc 1080
 QY 1107 cagcgaacacatcttctcgaagaatgccacttgtgtgagctgagccgaagatataagttccc 1166
 Db 1081 cagcgaacacatcttctcgaagaatgccacttgtgtgagctgagccgaagatataagttccc 1140
 QY 1167 caattcccatgtcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcat 1226
 Db 1141 caattcccatgtcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcat 1200
 QY 1227 caagatgtcaagaacaaactatcaaggttccacttcttaacaaacc 1271
 Db 1201 caagatgtcaagaacaaactatcaaggttccacttcttaacaaacc 1245

RESULT 11
 AAH43075
 ID AAH43075 standard; DNM: 1245 BP.

AC AAH43075;
 DT 15-OCT-2001 (first entry)
 XX

DE Nucleotide sequence of a human TGR-1 protein.

KW TGR-1; neuromedin U; hypertension; stress disease; ss.

OS Homo sapiens.

EH Key Location/Qualifiers

FT CDS 1..1245

FT /product= "TGR-1"

FT /note= "no termination codon given"

PN MO200157524-A1.

PD 09-AUG-2001.

PF 02-FEB-2001; 2001WO-JP00746.

PR 04-FEB-2000; 2000JP-0032773.

PR 24-FEB-2000; 2000JP-0052252.

PR 30-MAR-2000; 2000JP-0097896.

PR 19-JUN-2000; 2000JP-0187536.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;

PI Okubo S;

Db 1165 acccccacagccctctcagtgacagatgtaacaagaacacatcaaacgtccact 1224
 QY 1261 ttacaaacctgatatcttcagagctgactccctc 1298
 Db 1225 ttacaaacctgatatcttcagagctgactccctc 1262
 RESULT 14
 AAD08024
 ID AAD08024 standard; cDNA; 1231 BP.
 AC AAD08024;
 DT 07-AUG-2001 (first entry)
 XX
 DE Rat G-protein coupled receptor, SNORF72 cDNA.
 XX
 KW Rat; G-protein coupled receptor; SNORF72; neuromedin U neuropeptide;
 KW NMU; inflammation; arthritis; autoimmune disease; septicemia; psychotic;
 KW mental retardation; transplant rejection; neurological disorder; anxiety;
 KW respiratory disorder; depression; schizophrenia; dementia; obesity; pain;
 KW gastrointestinal disorder; hypertension; hypotension; epilepsy; diabetes;
 KW ischaemia; stroke; cancer; sexual disorder; circadian disorder; anorexia;
 KW dermatological; psoriasis; Parkinson's disease; nausea; bulimia; allergy;
 KW Alzheimer's disease; AIDS; hormonal disorder; memory disorder; migraine;
 KW cardiovascular disorder; renal disorder; bone disease; delirium; asthma;
 KW Cushing's disease; dysmenorrhoea; antilanginal; cytostatic; osteoporosis;
 KW metabolic disorder; behavioural disorder; Addison's disease; dyskinesia;
 KW tranquiliser; antiulcer; antiaddictive; ss.
 XX
 OS Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 23..1210
 FT /tag= a
 FT /product= "Rat SNORF72 receptor"
 FT
 FT
 PN WO200144297-A1.
 PD 21-JUN-2001.
 XX
 PF 13-DEC-2000; 2000MO-US33787.
 XX
 PR 17-DEC-1999; 99US-0466435.
 PR 25-APR-2000; 2000US-0558099.
 PR 30-JUN-2000; 2000US-0609146.
 XX
 PA (SYNA-) SYNAPTIC PHARM CORP.
 PI Bonini JA, Lerman GS, Quan Y, Ogozalek K;
 XX
 DR MPI: 2001-390240/41.
 DR P-PSDB; AA03634.
 XX
 PT A purified mammalian SNORF62 or SNORF72 receptor protein for
 PT identification of compounds to treat e.g. inflammation, arthritis,
 PT autoimmune diseases, transplant rejection, AIDS, cancer -
 XX
 PS Claim 15; Fig 14; 256pp; English.
 XX
 CC The invention relates to human G-protein coupled receptors, SNORF62 and
 CC SNORF72 and their corresponding cDNA molecules. SNORF62 and SNORF72
 CC receptors are specific for neuromedin U (NMU) neuropeptides, hence they
 CC are also known as NMU receptors. The agonist and antagonist of NMU
 CC receptors are useful for treating an abnormality in a subject that is
 CC alleviated by decreasing or increasing the activity of NMU receptor.
 CC The NMU receptors serves as a valuable tool for designing drugs which are
 CC useful for treating various pathophysiological conditions such as
 CC inflammation, arthritis, autoimmune diseases, transplant rejection,
 CC graft vs host disease, bacterial, fungal, protozoan and viral infections,
 CC septicemia, AIDS, pain, psychotic and neurological disorders, including
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory

CC loss, epilepsy, neuromotor disorders, respiratory disorders, asthma,
 CC eating/body weight disorders including obesity, bulimia, diabetes,
 CC anorexia, nausea, hypertension, hypotension, vascular and cardiovascular
 CC disorders, ischaemia, stroke, cancers, sexual disorders, circadian
 CC disorders, renal disorders, bone diseases including osteoporosis, benign
 CC prostatic hyper trophy, gastrointestinal disorders, nasal congestion,
 CC dermatological disorders such as psoriasis, allergies, Parkinson's
 CC disease, Alzheimer's disease, acute heart failure, angina disorders,
 CC delirium and dyskinesias such as Huntington's disease. They can also be
 CC used to regulate steroid hormone disorders, epinephrine release
 CC disorders, electrolyte balance disorders, endocrine disorders, memory
 CC disorders, somatosensory disorders, metabolic disorders, behavioural
 CC disorders, drug addiction, migraine, Addison's disease, Cushing's
 CC disease, prevent miscarriage, induce labour or to treat dysmenorrhoea.
 CC The present cDNA sequence encodes rat G-protein coupled receptor,
 CC SNORF72 related to the invention.
 XX
 SQ Sequence 1231 BP; 253 A; 367 C; 292 G; 319 T; 0 other;
 XX
 Query Match 61.6%; Score 799.6; DB 22; Length 1231;
 Best Local Similarity 79.2%; Pred. No. 4,9e-182;
 Matches 968; Conservative 0; Mismatches 239; Indels 15; Gaps 1;
 QY 18 tgaatttaatgtcagagatggaataaactcagatgcttcctgataccagcaaga 77
 Db 5 tgaatttaatgtcagagatggaataaactcagatgcttcctgataccagcaaga 55
 QY 78 ctagaagatcattccagaaacacctgaaacagcagcagagatattctgcttcctgc 137
 Db 56 -----gactctctcagaaagacttgaaacagcagcagagatattctgcttcctgc 109
 QY 138 ggaacctgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 197
 Db 110 ggaacctgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 169
 QY 198 gtgggggtcatgtgcaatgtcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 237
 Db 170 gtgggggtcatgtgcaatgtcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 229
 QY 258 acgcccccaactactactctctcagcctgtgctgtgtgtgtgtgtgtgtgtgtgtgt 317
 Db 230 acaccccccaactactactctctcagcctgtgctgtgtgtgtgtgtgtgtgtgtgtgt 289
 QY 318 ggaatgcacctgtgagatgtgagatgtgagatgtgagatgtgagatgtgagatgtgagat 377
 Db 290 ggaatgcacctgtgagatgtgagatgtgagatgtgagatgtgagatgtgagatgtgagat 349
 QY 378 ggcgtactactcaagacgagcagcagcagcagcagcagcagcagcagcagcagcagc 437
 Db 350 ggcgtactactcaagacgagcagcagcagcagcagcagcagcagcagcagcagcagc 409
 QY 438 accacgctcagcgtgagagcagcagcagcagcagcagcagcagcagcagcagcagc 497
 Db 410 accacgctcagcgtgagagcagcagcagcagcagcagcagcagcagcagcagcagc 469
 QY 498 agcacccggcgccggcccgcccgcccgcccgcccgcccgcccgcccgcccgcccgcccg 557
 Db 470 agcacccggcgccggcccgcccgcccgcccgcccgcccgcccgcccgcccgcccgcccg 529
 QY 558 tccctgcccac 617
 Db 530 tccctgcccac 589
 QY 618 gtcccaaggtcgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 677
 Db 590 gtcccaaggtcgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 649
 QY 678 caagtcacccct 737
 Db 650 caagtcacccct 709
 QY 738 ctcatgcaactcagactaagaagaacacacacacacacacacacacacacacacacacac 797

Db	710	ctctctggtgctctcagtgctgaaagaaagtgtatctcccttgagcgcaaaagctggtctgtaat	769
Qy	798	attcaaaagaccctcgcagaaataatcagtaacaaagaatgctgttctgtcttctgacttatt	857
Db	770	attcaagaccctctcagaaagatcagtaacaaagaatgctgttctgtcttctgacttatt	829
Qy	858	gctatctgctgtgagcccgcttcccaatctgacgaactctcttcagtcttctgtggaagatg	917
Db	830	gccctctgcctgcagacccccccttccatctgagacgagctcttcttcagacttctgtgaaagatg	889
Qy	918	agttaaacctctgagctgtctgtgttcaaacccgctccatctgtgttcaagtgatctctaac	977
Db	890	acaaagacctctgtgctgtctgtgttcaaacctccatccatgctgtatctcagtgctctcttat	949
Qy	978	ctgagctcagctgctcaaaccccatatctataactactctgtctcgcagctctccagcgaa	1037
Db	950	ctgagctcgcagcgcaaaccccatatctataactctctgtctcgcagctctccagcgagcc	1009
Qy	1038	ttccagaagtgtatctctcttcccaaaacagctgagacatcccaagatgaaccacagt	1097
Db	1010	ttctcgaagaatgctgtctctcccttaccctgtgcaaatggtgtccatcccggaatctgcgaaga	1069
Qy	1098	ccacacttccagcgcaaacatctctcgtgaagaagtccacttctgtgagctgacgaagaat	1157
Db	1070	ctctccagcccaagaagatacatctctctgtacagaagtgtcaactctgtgagctgacgaagaat	1129
Qy	1158	ataagttcccaatcccatctcagtaactcaatcagtaactcaactcccaacagccctc	1217
Db	1130	gcagagcccccagatctccctgtctcagtaactcaatcacaacacacacacttaccacgccc	1189
Qy	1218	ctctagtgaacagatgtctcaaga	1239
Db	1190	ctgtcagagaaggtacacataa	1211

RESULT	15
AAH43078	
ID	AAH43078 standard; DNA; 1165 BP.
XX	
AC	AAH43078;
XX	
DT	15-OCT-2001 (first entry)
XX	
DE	Nucleotide sequence of a rat TGR-1 protein.
XX	
KW	TGR-1; neuromedin U; hypertension; stress disease; ss.
XX	
OS	Rattus sp.
XX	
FH	Key
FT	Location/Qualifiers
FT	1..1185
FT	/tag=a
FT	/product="TGR-1"
ET	/note="no termination codon given"
XX	
PN	MO200157524-A1.
PD	
XX	
PD	09-AUG-2001.
XX	
PP	02-FEB-2001; 2001WO-JP00746.
XX	
PR	04-FEB-2000; 2000JP-0032773.
PR	24-FEB-2000; 2000JP-0052252.
PR	30-MAR-2000; 2000JP-0097896.
PR	19-JUN-2000; 2000JP-0187536.
XX	
PA	(TAKE) TAKEDA CHEM IND LTD.
PI	Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H,
PI	Okubo S;
DR	WPI: 2001-488917/53.

DR P-PSDB; AAG63367.
XX
PT Identifying predicted or actual structures of two or more members of a
PT chemical or physical library by mass spectrometry comprising
PT correlating molecular mass measurements of two or more members with a
PT shared chemical history -
XX
PS
XX Claim 9; Page 91-92; 95pp; Japanese.
XX
CC The present sequence encodes a rat TGR-1 protein. The specification
CC describes a method of screening a compound, which is capable of bindin
CC properties of neuromedin U to TGR-1. The method is useful for screenin
CC preventatives and remedies for hypertension, stress diseases, etc..
CC TGR-1 antagonists are also useful for treating the same diseases.
XX
XX Sequence 1185 BP; 242 A; 360 C; 277 G; 306 T; 0 other;
SQ

Query Match	60.68;	Score 787.2;	DB 22;	Length 1185;
Best Local Similarity	79.38;	Pred. No. 4.5e-179;		
Matches 952; Conservative	0;	Mismatches 233;	Indels 15;	Gaps 1

QY	36	atggaaaaaacattcgaatcgtctcccgaggtcttaccacagcaaaactagaaatccattccag	95
Db	1	atggaaaaaacattcgaatcgtctcccgaggtcttaccacagcaaaactagaaatccattccag	45
QY	96	aaacaaccgcaaacagacacagagagatcctcgtccctctccctcgcacacccgcgcagccac	155
Db	46	aagtaacttgaacaagacagagagagatcctcgtccctctccctcgcacacccgcgcagctac	105
QY	156	ttctctcccccgcgtctcgtcggtgtatgtatgccaattcttggctgggggtcatctggcaat	215
Db	106	ctatccctctccggctgtctcgtcggtgtatgtatgccaattcttggctgggggtcatctggcaat	165
QY	216	gtccctgggtgtccctgttatctctgagcaacagcggtatagaagacgcccacacactatcc	275
Db	166	ctctcgggtgtccctgttatctctgagcaacacggtatagaagacgcccacacactatcc	225
QY	276	ctctctcaaccgcggcggtctctctgacccctcgtctccctcgtctctcgaaatgtccctctgagtc	335
Db	226	ctctctcaaccgcggcggtctctctgacccctcgtctccctcgtctctcgaaatgtccctctgagtc	285
QY	336	tatgagatctggtcggaactaaccccttctctgtctcggggccggctggtcgtactctcaagacg	395
Db	286	tacgagatctggtcggaactaaccccttctctgtctcggggccggctggtcgtactctcaagac	345
QY	396	gccctctcttgagaccgtgtgtctctgcctccctccatccctccagatataccaacggtcagcggtg	455
Db	346	gccctctcttgagaccgtgtgtctctgcctccctccatccctccagatataccaacggtcagcggtg	405
QY	456	cgtacagctggagacatctctacacccgtctccggcgccaaacgtgacagagaccggcgccgggccc	515
Db	406	cgtacagctggagacatctctacacccgtctccggcgccaaacgtgacagagaccggcgccgggccc	465
QY	516	ctcagagatcctctcgcacatcgtctgtgggtctgtctccgtgtctctctccctggcccaacaacagc	575
Db	466	ctcagagatcctctcgcacatcgtctgtgggtctgtctccgtgtctctctccctggcccaacaacagc	525
QY	576	atccatggtcatcaagtctccactactctcccccgaatgggtccctgtgtcccaagttcgggcacc	635
Db	526	atccatggtcatcaagtctccactactctcccccgaatgggtccctgtgtcccaagttcgggcacc	585
QY	636	gttaaggtgcatcaagcccatgtgtgatacttacaattcatcatcccaagtcacccctccctcta	695
Db	586	gttaaggtgcatcaagcccatgtgtgatacttacaattcatcatcccaagtcacccctccctcta	645
QY	696	ttctactcctctcccaatgactgtatctcaagtgtctcctactactacatccatggtacactagacta	755
Db	646	ttctactcctctcccaatgactgtatctcaagtgtctcctactactacatccatggtacactagacta	705
QY	756	aagaaagacaaatctctcttgaggcgagatgaagaggaatgtcaaatatctcaaaagacctcgaga	815
Db	706	aagaaagatgtgaatccctcttgaggcgagaaagagaggtgtgaatatatctcaaaagacctctaga	765

PF 13-DEC-2000; 2000MO-US33787.
 XX
 PR 17-DEC-1999; 99US-0466435.
 PR 25-APR-2000; 2000US-0558099.
 PR 30-JUN-2000; 2000US-0609146.
 XX
 PA (SYNA-) SYNAPTRIC PHARM CORP.
 XX
 PI Bonloli JA, Lerman GS, Quan Y, Ogozalek K;
 XX
 DR WPI; 2001-390240/41.
 DR P-PDB; AA03634.
 XX
 PT A purified mammalian SNORF62 or SNORF72 receptor protein for
 PT identification of compounds to treat e.g. inflammation, arthritis,
 PT autoimmune diseases, transplant rejection, AIDS, cancer -
 XX
 PS Claim 15; Fig 14; 256pp; English.

XX The invention relates to human G-protein coupled receptors, SNORF62 and
 CC SNORF72 and their corresponding cDNA molecules. SNORF62 and SNORF72
 CC receptors are specific for neuromedin U (NMU) neuropeptides, hence they
 CC are also known as NMU receptors. The agonist and antagonist of NMU
 CC receptors are useful for treating an abnormality in a subject that is
 CC alleviated by decreasing or increasing the activity of NMU receptor.
 CC The NMU receptor serves as a valuable tool for designing drugs which are
 CC useful for treating various pathophysiological conditions such as
 CC inflammation, arthritis, autoimmune diseases, transplant rejection,
 CC graft vs host disease, bacterial, fungal, protozoan and viral infections,
 CC septicemia, AIDS, pain, psychotic and neurological disorders, including
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, neuromotor disorders, respiratory disorders, asthma,
 CC eating/body weight disorders including obesity, bulimia, diabetes,
 CC anorexia, nausea, hypertension, hypotension, vascular and cardiovascular
 CC disorders, ischemia, stroke, cancers, sexual disorders, circadian
 CC disorders, renal disorders, bone diseases including osteoporosis, benign
 CC prostatic hypertrophy, gastrointestinal disorders, nasal congestion,
 CC dermatological disorders such as psoriasis, allergies, Parkinson's
 CC disease, Alzheimer's disease, acute heart failure, angina disorders,
 CC delirium and dyskinesias such as Huntington's disease. They can also be
 CC used to regulate steroid hormone disorders, epinephrine release
 CC disorders, electrolyte balance disorders, endocrine disorders, memory
 CC disorders, somatosensory disorders, metabolic disorders, behavioral
 CC disorders, drug addiction, migraine, Addison's disease, Cushing's
 CC disease, prevent miscarriage, induce labour or to treat dysmenorrhoea.
 CC The present cDNA sequence encodes rat G-protein coupled receptor.
 CC SNORF72 related to the invention.
 CC
 XX
 SQ Sequence 1231 BP; 253 A; 367 C; 292 G; 319 T; 0 other;

Query Match 100.0%; Score 1231; DB 22; Length 1231;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gttgtgattttaagctcagtaattggaaaaacttgaaatgcttcttgatccagatcc 60
 DB 1 gttgtgattttaagctcagtaattggaaaaacttgaaatgcttcttgatccagatcc 60
 QY 61 tctcagaagtaacttgaaacagcacagagagatcttgagccacactgtgagcccaagcg 120
 DB 61 tctcagaagtaacttgaaacagcacagagagatcttgagccacactgtgagcccaagcg 120
 QY 121 cagtgacataccctcccggtgctctgagcctatgagcctatcttcttggtgggtaatt 180
 DB 121 cagtgacataccctcccggtgctctgagcctatgagcctatcttcttggtgggtaatt 180
 QY 181 gggcaatctctctggtgcaaggtaattgtcgcacatacgaacttgaagacaccaccaa 240
 DB 181 gggcaatctctctggtgcaaggtaattgtcgcacatacgaacttgaagacaccaccaa 240
 QY 241 ctactatctcttaagcttgagatcgaagatcagctgctgctcttgaggatgcctct 300
 DB 241 ctactatctcttaagcttgagatcgaagatcagctgctgctcttgaggatgcctct 300

DB 241 ctactatctcttaagcttgagatcgaagatcagctgctgctcttgaggatgcctct 300
 QY 301 ggaatcttaagatgtggaacaattacccttccgttcggagccgttgagatccactt 360
 DB 301 ggaatcttaagatgtggaacaattacccttccgttcggagccgttgagatccactt 360
 QY 361 caagacagccctctcgaagactgtgcttctgctccatctcaagtccacaggttag 420
 DB 361 caagacagccctctcgaagactgtgcttctgctccatctcaagtccacaggttag 420
 QY 421 cgtagagcgtatgtggtcatttccaccccttcgagacagctggaagacagcgcg 480
 DB 421 cgtagagcgtatgtggtcatttccaccccttcgagacagctggaagacagcgcg 480
 QY 481 acgggacctagatctcagcctagcttggaactctctcgtgtgttcttcttgccaa 540
 DB 481 acgggacctagatctcagcctagcttggaactctctcgtgtgttcttcttgccaa 540
 QY 541 taccagatccatgagatccagatccagacattcccaacgaggtccctcgtacctgctc 600
 DB 541 taccagatccatgagatccagatccagacattcccaacgaggtccctcgtacctgctc 600
 QY 601 agccacctgcacagtgacaacacccatgtggtgtatgaactgtatcatccaagctacag 660
 DB 601 agccacctgcacagtgacaacacccatgtggtgtatgaactgtatcatccaagctacag 660
 QY 661 ctctctcttatactctcccaatgagccctcaccagcgtctcttatacttcttgaggct 720
 DB 661 ctctctcttatactctcccaatgagccctcaccagcgtctcttatacttcttgaggct 720
 QY 721 caggtctgaagagatgaaatcccttgagcgcaaaagtgtgctgtgaatatcatcacagac 780
 DB 721 caggtctgaagagatgaaatcccttgagcgcaaaagtgtgctgtgaatatcatcacagac 780
 QY 781 ctctagaagtcagtcacccaagtgctgttctgtctgtctgtctgttcttgacatctgcg 840
 DB 781 ctctagaagtcagtcacccaagtgctgttctgtctgtctgttcttgacatctgcg 840
 QY 841 gaccccttcacagtgtagccggtctctctcaccactctgtggaagatgtagacagtcctc 900
 DB 841 gaccccttcacagtgtagccggtctctctcaccactctgtggaagatgtagacagtcctc 900
 QY 901 ggtctgctgtgttcaacctcatcctcagtgatcaagtgatcttcttattctgaagctcgcg 960
 DB 901 ggtctgctgtgttcaacctcatcctcagtgatcaagtgatcttcttattctgaagctcgcg 960
 QY 961 ggtcaaccccatattatctataacctctgtctgcggcttcggcggtcccttcgaagt 1020
 DB 961 ggtcaaccccatattatctataacctctgtctgcggcttcggcggtcccttcgaagt 1020
 QY 1021 tgtctccctacagctgcaaatgtgtgcatacccgagatccggaacagagacctccagccca 1080
 DB 1021 tgtctccctacagctgcaaatgtgtgcatacccgagatccggaacagagacctccagccca 1080
 QY 1081 gaagatcatcttcttgacagaatgtcaacctgtgagactgtacagagatgtcagagcccca 1140
 DB 1081 gaagatcatcttcttgacagaatgtcaacctgtgagactgtacagagatgtcagagcccca 1140
 QY 1141 gttccctgtgctgagatcatcctcacaacacacacacttaccacggtccctgtgcagagaga 1200
 DB 1141 gttccctgtgctgagatcatcctcacaacacacacacttaccacggtccctgtgcagagaga 1200
 QY 1201 ggtacataaagagtggtcagaagagctc 1231
 DB 1201 ggtacataaagagtggtcagaagagctc 1231

RESULT 2
 AAH43078
 ID AAH43078 standard; DNA; 1185 BP.
 XX
 AC AAH43078;
 XX

DT 15-OCR-2001 (first entry)
XX Nucleotide sequence of a rat TGR-1 protein.
DE TGR-1; neuromedin U; hypertension; stress disease; ss.
KM Rattus sp.
OS
XX
FH Key Location/Qualifiers
FT 1..1185
FT /*tag= a
FT /product= "TGR-1"
FT /note= "no termination codon given"
XX
XX WO200157524-A1.
XX
XX 09-AUG-2001.
XX
XX 02-FEB-2001; 2001WO-JP00746.
XX
XX 04-FEB-2000; 2000JP-0032773.
PR 24-FEB-2000; 2000JP-0052252.
PR 30-MAR-2000; 2000JP-0097896.
PR 19-JUN-2000; 2000JP-0187536.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;
P1 Okubo S;
XX
DR WPI: 2001-488917/53.
DR P-PSDB: AAG63367.
XX
XX Identifying predicted or actual structures of two or more members of a
PT chemical or physical library by mass spectrometry comprising
PT correlating molecular mass measurements of two or more members with a
PT shared chemical history -
XX
XX Claim 9; Page 91-92; 95pp; Japanese.
XX
XX The present sequence encodes a rat TGR-1 protein. The specification
CC describes a method of screening a compound, which is capable of binding
CC properties of neuromedin U to TGR-1. The method is useful for screening
CC preventative and remedies for hypertension, stress diseases, etc..
CC TGR-1 antagonists are also useful for treating the same diseases.
XX
XX Sequence 1185 BP; 242 A; 360 C; 277 G; 306 T; 0 other:
SO

Query Match 95.6%; Score 1177; DB 22; Length 1185;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1180; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 23 atgggaaaaacttgaaatgctctcctgataccagatctctcctcagatgaagtacttgaaacagc 82
DB 1 atgggaaaaacttgaaatgctctcctgataccagatctctcctcagatgaagtacttgaaacagc 60
OY 83 acagaggagacttggcccaactgtgtgagacccaagcgagtgacttaacccttcggtg 142
DB 61 acagaggagacttggcccaactgtgtgagacccaagcgagtgacttaacccttcggtg 120
OY 143 tctgagccatgagctgatactctcctgctggtggaatgaggaactcttggtgagcag 202
DB 121 tctgagccatgagctgatactctcctgctggtggaatgaggaactcttggtgagcag 180
OY 203 gtgattgtcgaacatcagaacttgaagacacccaactactatctcttcagcttgaca 262
DB 181 gtgattgtcgaacatcagaacttgaagacacccaactactatctcttcagcttgaca 240
OY 263 ggtcgaagactctgtgctgtctgtggtgagatgagctcttggaatctcagagtgagcac 322
DB 241 ggtcgaagactctgtgctgtctgtggtgagatgagctcttggaatctcagagtgagcac 300

OY 323 aatacctcttcctgttcggccctgtggatgacttaactcaagacagccctctcgagact 382
DB 301 aatacctcttcctgttcggccctgtggatgacttaactcaagacagccctctcgagact 360
OY 383 gtgtgcttgcctccatctcagtgatgacacaggtttagcglagagcgtatgtggcatt 442
DB 361 gtgtgcttgcctccatctcagtgatgacacaggtttagcglagagcgtatgtggcatt 420
OY 443 gtccaccccttcctcgaagcgaagctggaagacagcggtgagggcccttaagatcttcagc 502
DB 421 gtccaccccttcctcgaagcgaagctggaagacagcggtgagggcccttaagatcttcagc 480
OY 503 ctgagctgagactctcctgtgtctcttctcttgcccaatacagatcagagtcagcag 562
DB 481 ctgagctgagactctcctgtgtctcttctcttgcccaatacagatcagagtcagcag 540
OY 563 ttccagacattcccaagcgggtcctcgttacctgtgctcagccaactgtacagtcaccaa 622
DB 541 ttccagacattcccaagcgggtcctcgttacctgtgctcagccaactgtacagtcaccaa 600
OY 623 cccatgtgtgtgataacttgatcatcacaagctaccagcttctcttctacatctctcca 682
DB 601 cccatgtgtgtgataacttgatcatcacaagctaccagcttctcttctacatctctcca 660
OY 683 atgacccctcatcagctcctctactactcctcatggtggtcgaagctgaagagatgaatcc 742
DB 661 atgacccctcatcagctcctctactactcctcatggtggtcgaagctgaagagatgaatcc 720
OY 743 ctgtaggcgaacaaatgtgctgtgaataatcacaagaccccttaagaagtcatgacccaag 802
DB 721 ctgtaggcgaacaaatgtgctgtgaataatcacaagaccccttaagaagtcatgacccaag 780
OY 803 atgctgttctgtctgtctcctgtgtgttgcacatctgtgagaccccttcacatgtgacagg 862
DB 781 atgctgttctgtctgtctcctgtgtgttgcacatctgtgagaccccttcacatgtgacagg 840
OY 863 ctctctcctcagcttctgtgaagagtgagacagatcctcctgctgtgtgttcaaccctc 922
DB 841 ctctctcctcagcttctgtgaagagtgagacagatcctcctgctgtgtgttcaaccctc 900
OY 923 catgtgtatcagatgtgtctcttcttctatctgagctcgcgtgacccaactatctataac 982
DB 901 catgtgtatcagatgtgtctcttcttctatctgagctcgcgtgacccaactatctataac 960
OY 983 ctccgctcgtcggcttcctcggcgccttctggaatgtgtctcctccactctgcaaatgg 1042
DB 961 ctccgctcgtcggcttcctcggcgccttctggaatgtgtctcctccactctgcaaatgg 1020
OY 1043 tgcacatcccgagcatcggccacaagagacctccagcccaagatcatctcttgcagaaa 1102
DB 1021 tgcacatcccgagcatcagcaagagagacctccagcccaagatcatctcttgcagaaa 1080
OY 1103 tgtcacctcgttgagatgtgacagaagatgacagggccccaagcttccctgtgcatcatc 1162
DB 1081 tgtcacctcgttgagatgtgacagaagatgacagggccccaagcttccctgtgcatcatc 1140
OY 1163 cacaacacacacacttaccagagcccccgtgtgagagaaggtacaa 1207
DB 1141 cacaacacacacacttaccagagcccccgtgtgagagaaggtacaa 1185

RESULT 3
AAD08008
ID AAD08008 standard; cDNA; 1298 BP.
XX
XX AAD08008;
XX
XX 07-AUG-2001 (first entry)
XX
XX Human G-protein coupled receptor, SNORF72 cDNA.
DE
XX
XX Human; G-protein coupled receptor; SNORF72; neuromedin U neuropeptide;
KW NMU; inflammation; arthritis; autoimmune disease; septicemia; psychotic;

KM mental retardation; transplant rejection; neurological disorder; anxiety
 KM respiratory disorder; depression; schizophrenia; dementia; obesity; pain
 KM gastrointestinal disorder; hypertension; hypotension; epilepsy; diabetes
 KM ischaemia; stroke; cancer; sexual disorder; circadian disorder; anorexia
 KM dermatological; psoriasis; Parkinson's disease; nausea; bulimia; allergy
 KM Alzheimer's disease; AIDS; hormonal disorder; memory disorder; migraine
 KM cardiovascular disorder; renal disorder; bone disease; delirium; asthma
 KM Cushing's disease; dysmenorrhoea; antitumoral; cytostatic; osteoporosis;
 KM metabolic disorder; behavioural disorder; Addison's disease; dyskinesia;
 KM tranquiliser; antitumor; antidiabetic; ss.

OS Homo sapiens.

	Key	Location/Qualifiers
FH	CDS	27.1274
FT		

PN WO200144297-A1.

PD 21-JUN-2001.

PF 13-DEC-2000; 2000WO-US33787.

PR 17-DEC-1999; 99US-0466435.

PR 30-JUN-2000; 2000US-0609146.

PA (SYNA-) SYNAPTIC PHARM CORP.

PI Bonlnl JA, Lerman GS, Quar

DR WPI; 2001-390240/41.

XX

PT Identification of

PS Claim 42; Fig 3; 256pp; English.

CC The Invention relates to human c

CC receptors are specific for neuromedin U (NMU) neuropeptides, hence they
CC are also known as NMU receptors. The agonist and antagonist of NMU
CC receptors are useful for treating an abnormality in a subject that is
CC alleviated by decreasing or increasing the activity of NMU receptor.
CC The NMU receptors serves as a valuable tool for designing drugs which are
CC useful for treating various pathophysiological conditions such as
CC inflammation, arthritis, autoimmune diseases, transplant rejection,
CC graft vs host disease, bacterial, fungal, protozoan and viral infections,
CC septicemia, AIDS, pain, psychotic and neurological disorders, including
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, neuromotor disorders, respiratory disorders, asthma,
CC eating/body weight disorders including obesity, bulimia, diabetes,
CC anorexia, nausea, hypertension, hypotension, vascular and cardiovascular
CC disorders, ischemia, stroke, cancers, sexual disorders, circadian
CC disorders, renal disorders, bone diseases including osteoporosis, benign
CC prostatic hypertrophy, gastrointestinal disorders, nasal congestion,
CC dermatological disorders such as psoriasis, allergies, Parkinson's
CC disease, Alzheimer's disease, acute heart failure, angina disorders,
CC delirium and dyskinesias such as Huntington's disease. They can also be
CC used to regulate steroid hormone disorders, epinephrine release
CC disorders, electrolyte balance disorders, endocrine disorders, memory
CC disorders, somatosensory disorders, metabolic disorders, behavioural
CC disorders, drug addiction, migraine, Addison's disease, Cushing's
CC disease, prevent miscarriage, induce labour or to treat dysmenorrhea.
CC The present cDNA sequence encodes human G-protein coupled receptor,
CC SNORF72.

SQ Sequence 1298 BP; 278 A; 405 C; 282 G; 333 T; 0 other;

Query Match	65.08;	Score 799.6;	DB 22;	Length 1298;
Best Local Similarity	-79.28;	Pred. No. 3.2e-227;		
Matches 968;	Conservative 0;	Mismatches 239;	Indels 15;	Gaps 1;

QY	5	tgagtttaagctcagaaagaaagaaatgaaatgctcttcctgataccac-----	55
Db	18	tgagattttaaagtaagtaagaaacattcagaaatgctcttcctgataccacagagaa	77
QY	56	-----gatcctctccatagaaagtaactgaaacagacacagagaaagtaacttggccacactgtgt	109
Db	78	ctagaagaatccatctccagaaacacacgtgaaacagaccagagagtatactcgtctctctctgc	137
QY	110	ggaccccaagcgagctgacgtccatcccttcgcggtgtctgtgtgacctagcgctgatactctctg	169
Db	138	ggacacttcgcgagacccactctctctcccccgtctcgtgtgtatgtatgtgccaattttttgtg	197
QY	170	gttggggtaatgtggccaactctctcgtgtgtgtatgtatgtctgcgcacacagacttgaag	229
Db	198	gttggggtaatgtggccaatgtctcgtgtgtgtcgtgcgtgtatctctgcagacacagagactgaag	257
QY	230	acaccccaaccaactaactatctctctcaagcttgcagctcagatccatcgtgcgtccctgctctg	289
Db	258	acgcaccaaccaactactactccctctccagcgtgcgtctgcagctccctctgctctctctcct	317
QY	290	gggtgtcctctggaatactcagatggtgtggacaaattacacctcttcctgtttcgtggcgtctg	349
Db	318	ggaaatgccccctgggggtgctatgaaatggtgtgacaaataccctctctcgtgtccggcccg	377
QY	350	ggaagtactactcaagaaagacgccccctctccagagactgtgtcttgcctctcatctcaatgtgtc	409
Db	378	ggctgctactactcaagagcgtccctctgtgagacggtgtgtctgcctccatccctcaagatc	437
QY	410	acacacggttagctgtatgagcgctatgtgtgcacatgtgtccacaccttccggagccaagctgtag	469
Db	438	acccacccctcagcgtgtgagtcgtctatgttgcacatccactacacccgttccgcgccaactgtag	497
QY	470	agcacgcggtcgacggtggccccacagataccctccagacgtatgtctgagctctctctgtgtctt	529
Db	498	agcacccggtcggtcggtccctccagataccctccgatactgtcttgggtgtctctccgtgtctctc	557
QY	530	tcttggccaatatcacagataccatcgaatgttccagacacttcccaacgggtctccctcc	589
Db	558	tcccgtgccaacaacagacatccatccatgataagttcccaatactccccaattggtgccccgt	617
QY	590	gtacactgtgtcagacacactctgcagatgcacaaacccaaatgttgggtgtataaacttgaatc	649
Db	618	gtcccagatgttcggcacactgtgacgtgtacataaagcccacgtgtgataccaaattccatctcc	677
QY	650	caagctacacagcttccctctctctacatccctcccaatgacccctacagcttccctctacac	709
Db	678	cagtttaacctctctctctctctactaactccctcccaatgacgtatcaatgttctctactac	737
QY	710	ctcatgtgtgtctcaggtctgtgaagagagatgtaattccctctgtggtcgaaacaagtgtgctgtgaat	769
Db	738	ctctatgtgcactcagacataaagaagaagacaaatctctctgtgagcgatgtgaagtggaattgaaat	797
QY	770	attcaacgaacctctagaagaagtacatgcatacaagaatgtctgttgcctgtgtccgtgtgtt	829
Db	798	attcaaaagacccctcgaaaataatcagttcaaaaagaaatgtgtgtgtctgtgtcctagtgatc	857
QY	830	ggcaactctctgnaacccctctccatctgtgnaacgggtcttctctacgctttgtgtgnaagatgtg	889
Db	858	gctatctgtgtgtggcccggttccaaatgacagcgtctctctctcaagcttgtgtggaagtggtg	917
QY	890	acagagttccctgtgctgtgtgttcaaacctcaatccatgtgtgtatcaagtggtctctcttata	949
Db	918	agtataatccctgtgtgtgtgttcaaacctcgttcaacatgtgtgtgtgtgtgtgtctctctac	977
QY	950	cttgggtctccgggtgcacaccccatatctataaactctcttccggcggtctccggcgccgc	1009
Db	978	cttgggtctccagctgtgcacaccccatatctataaactcactgtctccggcggtctccagagcga	1037
QY	1010	tttcgnaatgtgtctccctccatctgnaaatgtgtgcatctcccggtgatactgcgcaagga	1069

Db 1038 ttccagatgcatctctcttccacaacagtgccactcccaagacagaccacagtg 1097
Oy 1070 cctccagccaggaatcatctcttgacagaatgcactctgtagagctgacagagat 1129
Db 1098 ccactgcgcagcgaatacatctctctgacagatgccactctgtgagctgacagagat 1157
Oy 1130 gcagcccccagctccctcgtgacatccatccacaacacacttaccagggcccc 1189
Db 1158 atagtcaccaatcccatctgacatccatgcacaactccactcccaacagccctc 1217
Oy 1190 tctgcagagaggtacacaa 1211
Db 1218 tctagtcacagatgtcaagaa 1239

RESULT 4
AAF80322
ID AAF80322 standard: DNA; 1658 BP.
XX
AC AAF80322:
XX
DT 29-JUN-2001 (first entry)
XX
DE Nucleotide sequence of a G-protein coupled receptor IGS4 long version.
XX
KW Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
KW nervous system disorder; psychiatric disorder; Parkinson's disease;
KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW gynecological disorder; ss.
XX
KW Homo sapiens.
XX
OS
FH Key Location/Qualifiers
FT 55..1302
FT CDS /*tag=a
FT /product="G-protein coupled receptor IGS4A long version"
XX
XX MO200125269-A2.
XX
XX PD 12-APR-2001.
XX
XX PF 25-SEP-2000; 2000WO-EP09584.
XX
XX PR 24-SEP-1999; 99EP-0203140.
XX PR 24-SEP-1999; 99NL-1013140.
XX PR 28-JUL-2000; 2000EP-0202683.
XX PR 31-JUL-2000; 2000US-0222047.
XX
XX PA (SOLV) SOLVAY PHARM BV.
XX
XX PI Deleersnyder W, Berger C, Loeken C, Nys G, Venema J;
XX WPI: 2001-273568/28.
XX DR P-PSDB; AAB67802.
XX
XX PT New G-protein coupled receptors and the polynucleotides encoding them,
XX PT useful for preventing, ameliorating or correcting nervous system
XX PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
XX PT or cancers -
XX
XX PS Claim 1; Page 75-77; 102pp; English.
XX
XX CC The present sequence encodes the long version of a human G-protein
XX CC coupled receptor designated IGS4A. IGS4 exists in two polymorphic forms,
XX CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
XX CC for preventing, ameliorating or correcting dysfunction or diseases.
XX CC These diseases include peripheral nervous system, psychiatric and central
XX CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal

CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
XX Sequence 1658 BP; 377 A; 490 C; 346 G; 445 T; 0 other:
SQ

Query Match 65.0%; Score 799.6; DB 22; Length 1658;
Best Local Similarity 79.2%; Pred. No. 3.6e-227;
Matches 968; Conservative 0; Mismatches 239; Indels 15; Gaps 1;

Oy 5 tggatttaagctcagtaatgggaaacttgaataatgtcttcctgataccac----- 55
Db 46 tggatttaatgtcagagatggaataaacttcgaatgtcttcctgataccacagagaa 105
Oy 56 -----gattcctcatgaagtacttgaacagcacagagggagtacttggccactgtgt 109
Db 106 ctagaagaacccatccccaagaacacttgaacagcacagagagatcttgcctcttcgc 165
Oy 110 ggaaccaagcagagtgacatcccttcctcggtctgtgagcctatgcctgtactctctg 169
Db 166 ggaactcgcgagacacatctctctccctccggtctgtgtgtgtatgtatgtgcaatttgg 225
Oy 170 gtgggggttaatgggaatctctgtgtgtgtatgtatgtgttcgcgaatcgaacttgaag 229
Db 226 gtgggggttattgtgaatgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 285
Oy 230 aacccacaactactatctctcagctgtgagctgtgagctgtgtgtgtgtgtgtgtgtgt 289
Db 286 agccccaacactactactctcagctgtgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 345
Oy 290 ggaatgcctcttgaatcactacagatgtgtgacaaatlaaccttctctgtctgggctgtg 349
Db 346 ggaatgcctcttgaatcactacagatgtgtgacaaatlaaccttctctgtctgggctgtg 405
Oy 350 ggaatgcctcttgaatcactacagatgtgtgacaaatlaaccttctctgtctgggctgtg 409
Db 406 ggaatgcctcttgaatcactacagatgtgtgacaaatlaaccttctctgtctgggctgtg 465
Oy 410 accaaggttaggttagagctatgtgtgacatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 469
Db 466 accaaggttaggttagagctatgtgtgacatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 525
Oy 470 agcagcgagcagcgccctcagatccctcagacagctagctgtgtgtgtgtgtgtgtgtgt 529
Db 526 agcagcgagcagcgccctcagatccctcagacagctagctgtgtgtgtgtgtgtgtgtgt 585
Oy 530 tcttggccaatacagatcactatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 589
Db 586 tccctgccaacacagatcactatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 645
Oy 590 gtaactgtgtgagcagcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 649
Db 646 gtccaggttgcgcacccgt 705
Oy 650 caagtacagcttctctctcactatccctcccaatgacctatcagctgtgtgtgtgtgtgt 709
Db 706 caggtacactctctctatctcactctcccaatgacctatcagctgtgtgtgtgtgtgtgt 765
Oy 710 ctcatgtgtgtcagctgtgaagatgtatccctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 769
Db 766 ctcatgtgtgtcagctgtgaagatgtatccctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 825

Dh	526	agcaccccgvgcgcgccctccagaaatccctcgtgcattgctcgggctctctccgtgccttc	585
Qy	530	tccttcgccaatlaacagacatccatcgtgcataaagtctccagacattccccaaggtctcc	589
Dh	586	tcacctgccaacacacagacatccatcgtgcataaagtctccatctatctccccaatgggtccctg	645
Qy	590	gtacctgctcagaccacctgcagagtcaaccaaacatgctgggtgtataattgatactgatactc	649
Dh	646	gtcccaaggtctcggccacctgtacggtcatcaaaagcccaatctgtgcataattcatcatctc	705
Qy	650	caagctacacagctctctctctctacatctccctccaaatgcacccctcaatcagcgtccctactc	709
Dh	706	caggtcacacctctccctcatctctacacctctccctcccaatgcacacagctgtccctctacc	765
Qy	710	ctcacctcgggtctccaggtctgaaagagatgaaatccctctcagagcgaaacaaagtgcctgaaat	769
Dh	766	ctcatcgtgacatccagactcaaaagaaagaaacaaatctctctgagcgagatgaaaggaatgcacaat	825
Qy	770	attcaacagaccctctagaaagtacatcaccaagaatgctgttctgtctcgtccctgtgtt	829
Dh	826	attcaaaagacccttcagaaaatacagtaacaagaatgctgttctgtctcgtctagctgt	885
Qy	830	gacatctgctgagacccctctccatctgtagacaggctctctctccagcttctggagaggtg	889
Dh	886	gctatctgtctgggcccgtgtcccaattgcacaggaattctctctccagcttctggagaggtg	945
Qy	890	acaaagatccacgtgctgctgtctcaaacatcatccatgtgtgtacaaagtgcctcttcat	949
Dh	946	agtgcataatcccgctgctgtcgtgtctcaaacctcgtccatgtgtgtgcaggtgtctctctac	100
Qy	950	ctgagctcgcggtctcaaacaccatataataacctctcgtccggtcgtcccgcgcc	100
Dh	1006	ctgagctcaggtctgtaaacaccatataataacctctcgtccggtcgtcccgcgcc	106
Qy	1010	cttcgaaatgtctgtccctccacactgcagaaatggtgcataatcccggaatcagcagagga	106
Dh	1066	cttcgaaatgtgcatactctctctccacaacaagtgcataatcccggaatcagcagaggt	112
Qy	1070	ccctcagagccccaagaatcatctctctgcagaaagtcatccctcgtgcaggtcagacaggaat	112
Dh	1126	ccacactgcccagcggaacatactctctgcagaaagtccaacttctgtgcaggtcagacaggaat	118
Qy	1130	gcagagcccccaggtccctcgtctcagttcatcatccacaacaacaacacttaacagagccccc	118
Dh	1186	ataagttccccaattcccaatgcagtcatcatcatgacaacaatctcaatcccaacagccctc	124
Qy	1190	ctgtcagagagaggtacataaa	1211
Dh	1246	ctcatgtgaacagagatgtcaagaa	1267

RESULT	6	
AAF80324		
ID	AAF80324	standard; DNA; 1658 BP.
XX		
AC	AAF80324;	
XX		
DT	29-JUN-2001	(first entry)
XX		
DE	Nucleotide sequence of a G-protein coupled receptor IGS4B long version.	
XX		
XX	Human: G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia; nervous system disorder; psychiatric disorder; Parkinson's disease; episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke cardiovascular disease; heart failure; angina pectoris; obesity; emesis; motility disorder; myocardial infarction; hypertension; dyslipidemia; gastrointestinal disorder; inflammatory bowel disease; osteoporosis; inflammation; infection; pain; cancer; immune disorder; allergy; sepsis; gynecological disorder; ss.	
KW		
XX		
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers

```

FT      CDS      55..1302
FT      /*tag= a
FT      /product= "G-protein coupled receptor IG54B long version"
PX      MM      MO200125269-A2.
XX      PD      12-APR-2001.
XX      PF      25-SEP-2000; 2000MO-EP09584.
XX      PR      24-SEP-1999; 99EP-0203140.
XX      PR      24-SEP-1999; 99NL-1013140.
XX      PR      28-OUL-2000; 2000EP-0202683.
XX      PR      31-JUL-2000; 2000US-0222047.
XX      PA      (SOLV ) SOLVAX PHARM BV.
XX      PI      Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
XX      DR      WPI; 2001-273568/28.
XX      DR      P-PSDB; AAB67804.
XX      PT      New G-protein coupled receptors and the polynucleotides encoding them,
XX      PT      useful for preventing, ameliorating or correcting nervous system
XX      PT      disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
XX      PT      or cancers
XX      PS      Claim 1; Page 83-85; 102pp; English.
XX
XX      The present sequence encodes the long version of a human G-protein
XX      CC      coupled receptor designated IG54B. IG54 exists in two polymorphic forms,
XX      CC      IG54A and IG54B. The IG54 receptors and IG54 polynucleotides are useful
XX      CC      for preventing, ameliorating or correcting dysfunction or diseases.
XX      CC      These diseases include peripheral nervous system, psychiatric and central
XX      CC      nervous system disorders (e.g. schizophrenia, episodic paroxysmal
XX      CC      anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
XX      CC      stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
XX      CC      myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
XX      CC      gastrointestinal disorders (e.g. inflammatory bowel disease or motility
XX      CC      disorders), osteoporosis, inflammations, infections (e.g. bacterial,
XX      CC      fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
XX      CC      sepsis or gynecological disorders. Agonists or antagonists of IG54 are
XX      CC      effective with regard to disorders of the nervous system, including the
XX      CC      central and peripheral nervous systems, disorders of the gastrointestinal
XX      CC      system, cardiovascular system, skeletal muscle, thyroid, lung or
XX      CC      genitourinary system, or immunological disease. The IG54 polynucleotides
XX      CC      are useful as diagnostic reagents for detecting under-expression,
XX      CC      overexpression or altered expression of IG54.
XX      CC
XX      Sequence 1658 BP: 378 A; 489 C; 346 G; 445 T; 0 other:
XX
XX      Query Match      64.7%; Score 796.4; DB 22; Length 1658;
XX      Best Local Similarity 79.1%; Pred. No. 3,2e-226;
XX      Matches 966; Conservative 0; Mismatches 241; Indels 15; Gaps 1
XX
XX      5 tggatttaagctcgtatgaggaaaacttgtaaaagcttcctcgatccac----- 55
XX      ||||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      46 tggattttaatcctcagggatgaaaaaacttcgaaagcttcctcgatccagcagagaa 105
XX      ||||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      56 -----gacccctcctagagacttgaaacaacagagagagtaacttgccaccctgtg 109
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      106 ctagaagaatcattcctcagaaacacccgaaacagcagcagcagagatattcgtcctctctgc 165
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      110 gagccaaagcagtgaaactatccctccggtgctggtgacctatgycgtatcctctcg 169
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      166 gagactcgcgcagcaccattctctcccccgcgtctgtgtatctgycgcaatttttgg 225
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      170 gtgggggtatattggcaattcttggtgtgcaaggtgatgttcgcagatcaagacttgaag 229
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      226 gtgggggtatattggcaattcttggtgtgctcgtgtatcttcgcagcaggtatgaag 285
XX      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX      230 aacaccacaactatctctcgaagcttggaagcttcgaatctcgatctgctgctgccttg 289

```

```

DB 286 acgcccacacactactctctcagccgctcctcgtctcgtactcctcgtctccgtcctt 345
QY 290 gggatgctcctcgtgaatctcagagatgtgacaaatacctctcctgttggcgcttg 349
DB 346 ggaatgccccctggaggtctctatgagatgtgacgaactacccctctctgtcggcccg 405
QY 350 ggaatgctactcagaacacacccctcctcgtgagctgtgtccttccctcctcagtgctc 409
DB 406 ggcgtgactctcagaagcgcgcctccttcttgagacgtgtgtcctcctcctcctcagc 465
QY 410 accacggttagcgttagagcgtatgtgtgcaatgtccacccttccgagccaagctgag 469
DB 466 accacccgttagcgttagagcgtatgtgtgcaatgtccacccttccgagccaagctgag 525
QY 470 agcagcgagcgagggcgccctcagagatcctcagcgttagctcgtgagcttctcgtg 529
DB 526 agcagcgagcgagggcgccctcagagatcctcagcgttagctcgtgagcttctcgtg 585
QY 530 tcttgcacatatacagacatcctcagatcctcagcgttagctcgtgagcttctcgtg 589
DB 586 tccctgcacacacacacacacacacacacacacacacacacacacacacacacacac 645
QY 590 gtaactgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 649
DB 646 gtcacagctcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 705
QY 650 caagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 709
DB 706 caggtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 765
QY 710 ctcatgagcgtcagagcgtcagagagagatgaaatccctcgtgagcgagacaagatgag 769
DB 766 ctcatgagcgtcagagcgtcagagagagatgaaatccctcgtgagcgagatgagagag 825
QY 770 attcacagacccctcagagagatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 829
DB 826 attcacagacccctcagagagatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 885
QY 830 gcatctgctcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 889
DB 886 gcatctgctcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 945
QY 890 acagagatccctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 949
DB 946 acgaaatccctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1005
QY 950 ctgagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1009
DB 1006 ctgagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1065
QY 1010 ctgagagatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1069
DB 1066 ctgagagatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1125
QY 1070 cctcagacccagagagatcctcctcctcagagagatgtaacatcgtgagctgagagagat 1129
DB 1126 ccaactcgtcagagagatcctcctcctcagagagatgtaacatcgtgagctgagagagat 1185
QY 1130 gcaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1189
DB 1186 atgggtccccaatctctatgtcagtcacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1245
QY 1190 tctgagagagagagagagagagagagagagagagagagagagagagagagagagagag 1211
DB 1246 tctagtgaaacagatgtcagaaga 1267

```

```

XX 29-JUN-2001 (first entry)
DE Nucleotide sequence of G-protein coupled receptor IGS4B short version.
XX
KW Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
KW nervous system disorder; psychiatric disorder; Parkinson's disease;
KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW gynecological disorder; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 64..1302
FT /tag= a
FT /product= "G-protein coupled receptor IGS4B short
FT version"
PN WO200125269-A2.
PD 12-APR-2001.
PE 25-SEP-2000; 2000MO-EP09584.
PF 24-SEP-1999; 99EP-0203140.
PR 24-SEP-1999; 99NL-1013140.
PR 28-JUL-2000; 2000EP-0202683.
PR 31-JUL-2000; 2000US-0222047.
XX
PA (SOLV ) SOLVAY PHARM BV.
PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
XX WPI; 2001-273568/28.
XX P-PSDB; AAB67805.
DR
PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers .
XX
PS Claim 1: Page 87-89; 102pp; English.
XX
CC The present sequence encodes the short version of a human G-protein
CC coupled receptor designated IGS4B. IGS4 exists in two polymorphic forms,
CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
CC for preventing, ameliorating or correcting dysfunctions or diseases.
CC These diseases include peripheral nervous system, psychiatric and central
CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
SQ Sequence 1658 BP; 378 A; 489 C; 346 G; 445 T; 0 other:

```

```

RESULT 7
ID AAF80325 standard; DNA; 1658 BP.
XX
AC AAF80325;

```

```

Query Match 64.7%; Score 796.4; DB 22; Length 1658;
Best Local Similarity 79.1%; Pred. No. 3, 2e-226;
Matches 966; Conservative 0; Mismatches 241; Indels 15; Gaps 1;

```

QY	5	tgratatttaagctcaatgataatvgrggaataacttgaaatvgtcttcctgfatccac-----	55
Db	46	tgratatttaagctcaatgataatvgrggaataacttgaaatvgtcttcctgfatccac-----	105
QY	56	-----gattccttcacagaaagtactcttgaaacagaaagagatgactcttgagccactgtgt	109
Db	106	ctagaaagatccatctccagaaacacactgaaacagaccgagagatatactggtcttcctctg	165
QY	110	ggaccacaaagccaaatgacactacactccctccggtgtctctgtgacataagctgcgactccctg	169
Db	166	ggaccctcggtcgacagccactctctccctcccggtgtctctgtgacataagctgcgactccctg	225
QY	170	gtggggggtgaatvgrggaactctctctgtgtgtgacatgvtgacttgcgaacatacagacttgaaag	229
Db	226	gtggggggtgaatvgrggaactctctctgtgtgtgacatgvtgacttgcgaacatacagacttgaaag	285
QY	230	acacccacacaaactactatctcttaagctcttgagactccaaatctgctgtgtccctgctctg	289
Db	286	acgcccacacaaactactatctcttaagctcttgagactccaaatctgctgtgtccctgctctc	345
QY	290	ggagatgcctcgtgaatctagaaatgvtggtgcacaaataacacttcctcgttcgggacctgtg	349
Db	346	ggagatgcctcgtgaggtgtctatgaaatgvtggtgcacaaataacacttcctcgttcgggacctg	405
QY	350	ggatgctactctcaagagacagccctcttcgaaagactgtgtgctcttgctccatctccagtgct	409
Db	406	ggctgctactctcaagagagccctcttcgaaagacggtgtgtgcttcgctccatctccagatc	465
QY	410	acacaggtctagcgttagagcgtctatgtgtgccaatgtgtccaccccttcgcgaacagctggag	469
Db	466	acacacgcgtcagcgtgtgagcgtctatgtgtgcatacctcaacccggttcgcgcgaacagctgag	525
QY	470	agacacgcgcgcagcgcgcgcctcccaagaaatccctcaagccttaagctgtctgtgtgtgtct	529
Db	526	agacacgcgcgcgcgcgcgcgcctcccaagaaatccctcaagccttaagctgtctgtgtgtgtct	585
QY	530	tcctctgcccacaaatacagacatccatccatgacatgacttccagacacttcccaacggtgcctcc	589
Db	586	tcctctgcccacaaacacagacatccatccatgacatgacttccacacttcccaatvggtccctg	645
QY	550	gtacctgtgctcagccacactgtgcacagtcacacaaaccaaagtgtgggtgtataactgtgatatc	649
Db	646	gtcccaggtctcgcacacactgtgacgttcatacagcccaatgtgtgactataacttcatcatc	705
QY	650	caagactacacagcttcctccctctctataatccctcccaatgacccctcaatcagcgtccctactac	709
Db	706	caagactacacctctctctctctctctctctctccctcccaatgacgtctcaatgtgtccctctac	765
QY	710	ctcaatvgggtcagagcgtgaagagagatcgtatccctcttgaggcgaaacaaagtgtgtgtgaat	769
Db	766	ctcaatvgggtcagagcgtgaagagagatcgtatccctcttgaggcgaaagaaatgtgaat	825
QY	770	attcaacagaacctctagaaagatcagtcacacaaagatgtcgttgttgtctgtgtcctgtgtt	829
Db	826	attcaaaagaacctctagaaagatcagtcacacaaagatgtcgttgttgtctgtgtcctgtgtt	885
QY	830	gcatctgtgtgtgacccctctccatgtgtgagacgggtctcttcaagctcttgvggaaaggtgtg	889
Db	886	gcatctgtgtgtgacccctctccatgtgtgagacgggtctcttcaagctcttgvggaaaggtgtg	945
QY	890	acaaagatccctcgtgtcgtgtgttcaacactatccatgtgtgtatacaggtgtctcttcat	949
Db	946	actgaaatccctcgtgtcgtgtgttcaacactatccatgtgtgtatacaggtgtctcttctac	100
QY	950	ctgagctcgcgcgtgtcaaaccccatatcatataaacctctcgtctcgcgcgtcgcgcgcgcgc	1000
Db	1006	ctgagctcgcgcgtgtcaaaccccatatcatataaacctcgtctcgcgcgtcgcgcgcgcgcgc	106
QY	1010	ttctgaaatgtgtgtctccctcactctgtaaaatgtgtccaatcccggaatcgcgcacagga	106
Db	1066	ttctgaaatgtgtgtctcttcttccacaaacaggtgtgtgacccctcagaaatgtgacacaggtgt	112

Oy	1070	cctccagccccagaatcatccttcttgacagaagtgcacctggtgagctgtgacaagat	1129
Dd	1126	ccacttcgcccgaggaaacatcttccttgacaagaatgcccattgttggagctgaccgaagat	1185
Oy	1130	gagagcccccaattcccttgctacgtcacatccacacacccaactttaccacagggcccc	1189
Dd	1186	ataagttcccacattcctatgtcagctaccccgtagcacaactctaccctccacagccctc	1245
Oy	1190	tgtcagaggagaagtlaccataa	1211
Dd	1246	tctagtgacagatgtcaagaa	1267
 RESULT AAZ33297 standard; cDNA; 1360 BP.			
ID	AAZ33297		
XX	AAZ33297;		
AC	AAZ33297;		
XX			
DT	21-FEB-2000 (first entry)		
XX			
DE	Human neurotensin-like receptor encoding cDNA.		
XX			
KM	Human; neurotensin-like receptor; NLR: G-protein coupled receptor;		
KW	central nervous system; anesthesia; analgesia; ds.		
XX			
OS	Homo sapiens.		
PN	WO955732-A1.		
PN			
PD	04-NOV-1999.		
XX			
PF	15-APR-1999; 99WO-SE00598.		
XX			
PR	24-APR-1998; 98SE-0001455.		
XX			
PA	(ASTR-) ASTRA PHARMA INC.		
PA	(ASTR) ASTRA AB.		
PI	Ahmad S, Cao J, O'Donnell D, Walker P;		
DR	WPT: 2000-052803/04.		
DR	P-PBDB; AAI52992.		
XX			
PT	Novel neurotensin-like receptor, useful for identifying agents for		
PT	producing anesthesia or analgesia -		
XX			
PS	Claim 11; Fig 1; 46pp; English.		
XX			
CC	The present sequence encodes a human G-protein coupled receptor		
CC	neurotensin-like receptor (NLR). The NLR polynucleotide and protein		
CC	can be used to isolate compounds that bind, (antagonise or alter		
CC	the activity or expression of the NLR. The NLR is a G-protein coupled		
CC	receptor which is expressed in the central nervous system and shares		
CC	homology with human neurotensin receptor. The receptors can be used in		
CC	assays to identify agents for producing anesthesia and analgesia.		
XX			
SQ	Sequence 1360 BP; 293 A; 423 C; 294 G; 350 T; 0 other:		
 Query Match 64.3%; Score 791.6; DB 21; Length 1360; Best Local Similarity 78.8%; Pred. No. 7.8e-225; Matches 963; Conservative 0; Mismatches 244; Indels 15; Gaps 1;			
Oy	5	tggattttaagctcgtaatggtgaaaacttggaaatgcttctctgatcac-----	55
Dd	56	tggattttaatgtcaggatgtaaaaaacttcagaatgcttctctgatctaccagagaaa	115
Oy	56	-----gacccctatgagagtactctgaacagcagacagagagttactgcccaccgttgt	109
Dd	116	ctagaagaatcatcattccagaacaacactgaaacagacagagagttactgtgcttcttcgc	175
Oy	110	ggaccacaagcagatgacctatcccttcgcgggtgtctgtgacctatgctgcttctctgy	169

Dh	176	ggaactcggcgagcgaactcttccttcctcccgctgctcgtgtagtgcatactgttcgtg	235
Oy	170	gtcgggtgtaalcyggcgaactctcgtgtgtgtacgtatgtgtccgacatacagacttggag	229
Dh	236	gtgggggtgcatactgcgaatgtctcgtgtgtccgtgtgtatcttcgacagcaaggctagaag	295
Oy	230	acaacacacaaactatactctcttcacagcttggagctccgaatctgtgttccgtgccttg	289
Dh	236	acgccacacaaactatactcttcacagccctgcgcgtgtcccttcgaacctcgtgtccctctt	355
Oy	230	ggagatgcctctggaatactacagaaagtgtgcagacaatatacccttccctgtctgcggcctg	349
Dh	356	ggaaagccccctggaggtctatagaaagtgtgcgcacaatacccttctgttctgcggccgtg	415
Oy	350	ggatgtctatactcaagagaacgcccctcttcgaaactcgtgtgccttgcctcaatcttcagtgc	409
Dh	416	ggctgtctatactcaagagaagcccctctcttgaaagcgtgtgcttgcctccatccctcagatc	475
Oy	410	accaaggtctaaagctagaagcgtctaatgtgtgcacatgtgcaccccttccgaagcaagctggag	469
Dh	476	acaacacgtcagcgtgtggagcgtctaaagtgtgcacatccaaacccgttccgcgcacaactcgag	535
Oy	470	agcacgcgctgagcggcgccctccataggaacctccagccctagtctggaagctctcgtgtgcttt	529
Dh	536	agcaccccgctgcgcggccctccctcagaaatccctcgtgacatcgtctcgtggcttccctcgtcttc	595
Oy	530	tccttggcccaatacacaagacatccagatataagttccagacacttcccaagcggtccctcc	589
Dh	556	tcctctggcccaaacacacagacatcccatgtgcataaagtctcaactaacttcccacatgggtccctg	655
Oy	550	gtaccctgtgtcagccacacctgtgcacagtaccacaaccccaatgtgtgtgtatatactgtatcaltc	649
Dh	656	gtcccaaggtctcgcacacactgtgaagctatcaagcccaatgtgagatcttaacaattcatcaltc	715
Oy	650	caagagtaacagacttccctctctctataatccctcccaatgaaacctatacagctctctctctac	709
Dh	716	caggacacactccctctctatacttcaactccctcccaactgtctacatgcagtgtctccctctctac	775
Oy	710	ctcaatcgtggtctcagcgtcgaagagagaaatgcgaatccctcttbaagcgcgaacaagatgtgtgaat	769
Dh	776	ctcaatcgtgacatcagatacgaagaagaaatctcctctgaagcagaaatgaagaggaatgaat	835
Oy	770	attcaacagaacctctagaaagatcagttcacagaatgtcgttctgtctcgtgcctcgttgtt	829
Dh	836	attcaaaagaacctctcagaaaaatcagttcaaaagaatgtcgttctgtctgtctctatgtt	895
Oy	830	gcaactctcgtggaacccccctccagatgtgagacgcgtctctcttgaagcttttcttgaagaagtgtg	889
Dh	896	gctatctcgtgtggcccgcttcccaatctgaaacgcgaatctctctccagcctctgtgtggagagatgtg	955
Oy	890	acagagatccctcgtgcctgtgtgttcaacctatcccatgtgtgtatcaaggtgtctctcttlat	949
Dh	956	actgaaatccctcgtgcctgtgtgttcaaacctctgtccaatgtgtgtgcaggtgtcttatactaac	1011
Oy	950	ctgtagctctcgcgtgtccaaacccaatatactataaactcctgtctctcggcgtctccggcgcgcc	1007
Dh	1016	ctgtagctctagctgtctcaaaccccaatatactataaactaactgtctctcgcgtctccagcgtagaga	1071
Oy	1010	cttcgaaatgtctgtcccccacactggaacatgtgtccatcccccggagatccgcaacagggga	1066
Dh	1076	cttcagaaatgtgactcttcttcttccacaacaaatgtgcactcccaagatatacccaacagttg	1133
Oy	1070	ctctccagaccgaagaatcaatctctctgtacagaagatgttaacctcgtgtgtagctgtgacagagat	1123
Dh	1136	ccacactcccgacggaacaaacttctcctctcgaagaatgtccacatctgtgtgagctgtgacgaagt	1191
Oy	1130	gcagagccccccagcttccctcgtgtgtagtatccatccacaacaaacaaacttaccagcgcccc	1188
Dh	1196	ataagttccccaattcccaatgactcagttacgtacccgtgtgacaaacttcaactcccaacagccctc	1251
Oy	1190	ctgtcagagagaggtacataaa	1211

DB	1256	tctagtgaacagatgtcaagaa	1277
RESULT	9		
AAH43072			
ID	AAH43072	standard; DNA; 1245 BP.	
XX			
AC	AAH43072;		
XX			
DT	15-OCT-2001	(first entry)	
XX			
DE		Nucleotide sequence of a human TGR-1 protein.	
XX			
KW	TGR-1; neuromedin U; hypertension; stress disease; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..1245	
FT		/*tag= a	
FT		/product= "TGR-1"	
FT		/note= "no termination codon given"	
XX			
PN	WO200157524-A1.		
XX			
PD	09-AUG-2001.		
XX			
PF	02-FEB-2001; 2001WO-JP00746.		
XX			
PR	04-FEB-2000; 2000JP-0032773.		
PR	24-FEB-2000; 2000JP-0052252.		
PR	30-MAR-2000; 2000JP-0097896.		
PR	19-JUN-2000; 2000JP-0187536.		
XX			
PA	(TAKE) TAKEDA CHEM IND LTD.		
XX			
PI	Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;		
PI	Okubo S;		
XX			
DR	WPI; 2001-488917/53.		
DR	P-PSDB; AAG63353.		
XX			
PT	Identifying predicted or actual structures of two or more members of a		
PT	chemical or physical library by mass spectrometry comprising		
PT	correlating molecular mass measurements of two or more members with a		
PT	shared chemical history -		
XX			
PS	Claim 9; Page 79-80; 95pp; Japanese.		
XX			
CC	The present sequence encodes a human TGR-1 protein. The specification		
CC	describes a method of screening a compound, which is capable of binding		
CC	properties of neuromedin U to TGR-1. The method is useful for screening		
CC	prelives and remedies for hypertension, stress diseases, etc..		
CC	TGR-1 antagonists are also useful for treating the same diseases.		
XX			
XX	Sequence 1245 BP; 268 A; 393 C; 268 G; 316 T; 0 other;		
Query Match	64.2%;	Score 790.8;	DB 22; Length 1245;
Best Local Similarity	79.2%;	Pred. No. 1.3e-224;	
Matches 958; Conservative	0; Mismatches 237;	Indels 15; Gaps 1;	
QY	17	tcaagtaatgggaaacttgaanaatgcttcctcgtgacacac-----gatcct	61
DB	4	tcaaggatggaagaaacttcaagaatgcttcctcgtgacacacacagaactagaagatcca	63
QY	62	ctcaagaagtaacttgaacagcacagaggaagtaacttcgcccacaccttctgtgagcccaagcgc	121
DB	64	ttccagaacacacctgaacagcacgcaggaagatctcgtgccttcctctggtgagacactcgagc	123
QY	122	agtgacatacccttcgcggtctctgtgacctatgcgctgatcttcctctgtgtgggtaatg	181
DB	124	agccacttccttcctccggtctctgtgtgtatgtgccaatcttctgtgtgtggtgcatc	183

Oy	182	ggaacatctcggggtgacatgagtgatgtgttcgcagcatccagactttgaagacaacccacac	241
Dd	184	ggcaaaatgcccggggtggtgcgcgggtgaaattctgcagaaacacaggtcattgaaagacgcccacac	243
Oy	242	tactacatctctcaagctcvgagactcgcagatccatgcctctgtctgcctctctctgagatgtgcctc	301
Dd	244	tactacatctctcaagctcvgagactcgcagatccatgcctctgtctgcctctctctgagatgtgcctc	303
Oy	302	gaaattctacagagatggtgccaattacaccttccctgttctcgagcctggtggaatgtaattc	361
Dd	304	gaagttctactgagagatggtgccaattacaccttccctgttctcgagcctggtggaatgtaattc	363
Oy	362	aagagcaagccctctcgagactcgt	421
Dd	364	aagagcaagccctctcgagactcgt	423
Oy	422	gttagagcgtctagtgtgccaattgttcaaccttcccgagccaagctgtgagagacagcggcgga	481
Dd	424	gttagagcgtctagtgtgccaattgttcaaccttcccgagccaagctgtgagagacagcggcgga	483
Oy	482	cgggcccctcagagatccctccagctctagctctgagagttctctcgtgtgtgtgtgtgtgtgtgtgt	541
Dd	484	cgggcccctcagagatccctccagctctagctctgagagttctctcgtgtgtgtgtgtgtgtgtgtgt	543
Oy	542	accgagatcccatgtgacatcaaggtttccagcaatttcccaacgggtgctccctcgtaactgtgta	601
Dd	544	accgagatcccatgtgacatcaaggtttccagcaatttcccaacgggtgctccctcgtaactgtgta	603
Oy	602	gccaactcgcagatctaccacaaaccaaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	661
Dd	604	gccaactcgcagatctaccacaaaccaaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	663
Oy	662	tctcctctctacatccctcccaatgaaacctatagcgtctctacactcaactcaatgagacatc	721
Dd	664	tctcctctctacatccctcccaatgaaacctatagcgtctctacactcaactcaatgagacatc	723
Oy	722	aggctgaaagagagatgaaatccctctgagcggaaacaaggtgagctgtgtaatatctacagaccc	781
Dd	724	aggctgaaagagagatgaaatccctctgagcggaaacaaggtgagctgtgtaatatctacagaccc	783
Oy	782	tcctgaaagatcagatcacacaagaagtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	841
Dd	784	tcctgaaagatcagatcacacaagaagtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	843
Oy	842	accacctctcactgtgagacgcgactctctcttcagactctgtgtgaaagagtgagacaagctccct	901
Dd	844	accacctctcactgtgagacgcgactctctcttcagactctgtgtgaaagagtgagacaagctccct	903
Oy	902	gctgctgtgtgttcaaacctccatccatggtgataccaagtgctctctttatcttgatgcctcgcg	961
Dd	904	gctgctgtgtgttcaaacctccatccatggtgataccaagtgctctctttatcttgatgcctcgcg	963
Oy	962	gtcaaaccccatctactataaacctcgtctcgcggcgctccggcgctccagagacatctccagact	1021
Dd	964	gtcaaaccccatctactataaacctcgtctcgcggcgctccggcgctccagagacatctccagact	1023
Oy	1022	gtctccctctaccgtgcaaaatggtgcacatcccgcgcatctgcgacaagagacatctccagccag	1081
Dd	1024	gtctccctctaccgtgcaaaatggtgcacatcccgcgcatctgcgacaagagacatctccagccag	1083
Oy	1082	aagatctactctctgaaagaaatgttccctctgtgagactgtgacagagatgtcagagcccccag	1141
Dd	1084	aagatctactctctgaaagaaatgttccctctgtgagactgtgacagagatgtcagagcccccag	1143
Oy	1142	tctccctgtctgatactcatccatcaacaacaacacatcttaccagcgccccctgtgtcagagagag	1201
Dd	1144	tctccctgtctgatactcatccatcaacaacaacacatcttaccagcgccccctgtgtcagagagag	1203
Oy	1202	gtaccatataa 1211	
Dd	1204	agtataaata 1213	

RESULT	10	
ADAD01123	standard; cDNA; 1248 BP.	
ADAD01123		
ADAD01123		
02-NOV-2000	(first entry)	
Human orphan G protein-coupled receptor hrup6 cDNA.		
Human; orphan G protein-coupled receptor; GPCR; hrup6; drug screening; transmembrane receptor; signal cascade; ss.		
Homo sapiens.		
Key	Location/Qualifiers	
CDS	1..1248	
FT	/*tag= a	
FT	/product= "hrup6"	
FT	/note= "Human orphan G protein-coupled receptor"	
MO2000031258-A2.		
02-JUN-2000.		
13-OCT-1999;	99MO-US23687.	
20-NOV-1998;	98US-0109213.	
16-FEB-1999;	99US-0120416.	
26-FEB-1999;	99US-0121852.	
12-MAR-1999;	99US-0123946.	
12-MAR-1999;	99US-0123949.	
28-MAY-1999;	99US-0136456.	
28-MAY-1999;	99US-0136457.	
28-MAY-1999;	99US-0136459.	
28-MAY-1999;	99US-0136567.	
28-MAY-1999;	99US-0137127.	
29-JUN-1999;	99US-0141448.	
29-SEP-1999;	99US-0156555.	
29-SEP-1999;	99US-0156633.	
29-SEP-1999;	99US-0156634.	
01-OCT-1999;	99US-0157280.	
01-OCT-1999;	99US-0157281.	
01-OCT-1999;	99US-0157282.	
01-OCT-1999;	99US-0157293.	
12-OCT-1999;	99US-0157294.	
12-OCT-1999;	99US-0416760.	
12-OCT-1999;	99US-0417044.	
(AREN-) ARENA PHARM INC.		
Chen R, Dang HT, Liaw CW, Lin I;		
WPI: 2000-400068/34.		
P-PSDB: AAY71296.		
Novel human orphan G protein-coupled receptors and the encoding cDNAs for use in the identification of G protein-coupled receptor agonists -		
Claim 21; Page 56-57; 102pp; English.		
The present sequence is a cDNA encoding hrup6, an endogenous human orphan G protein-coupled receptor (GPCR). The full length hrup6 cDNA was cloned by RT-PCR using human thymus cDNA as template.		
The orphan GPCR of the invention, like all GPCRs has seven transmembrane alpha helices with an extracellular N-terminus and an intracellular C-terminus. However, no endogenous ligands has yet been identified for the proteins of the invention. The orphan GPCRs may be used in the identification of their endogenous ligands, and to screen potential GPCR agonists and antagonists for use as pharmaceutical agents. The proteins		

CC may also be used in the study of GPCR-mediated signalling cascades, and
CC to elucidate their precise role in normal and diseased human conditions.
CC Nucleic acid encoding human orphan GPCRs may be used for tissue
CC localisation expression analysis to provide information about their
CC function in healthy and pathological states.

XX Sequence 1248 BP; 269 A; 393 C; 269 G; 317 T; 0 other;

Query Match 64.2%; Score 790.8; DB 21; Length 1248;
Best Local Similarity 79.2%; Pred No. 1.3e-224;
Matches 958; Conservative 0; Mismatches 237; Indels 15; Gaps 1;

```

OY 17 tcaagtaatggaagaacttgaagaatgtctctctgattccac-----gatcct 61
DB 4 tcaaggatggaagaacttcaagaatgtctctctgattccacgaagaactagaagttcca 63
OY 62 ctcaatgaatcttgaacagcagcagaagaagtaatttggccaccctgtgtgacccaagcgc 121
DB 64 ttcagaagaacacgtgaacacagcagcagagatatttgcctctcctctgcggaacctgcgc 123
OY 122 agtgaactatcccttcctcgtgtctgtgacctatgacctatctctctgttggggtaatg 181
DB 124 agccaatcttccctccctcgtgtctgtgacctatgacctatctctctgttggggtaatg 183
OY 182 ggcacatctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 241
DB 184 ggcacatctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 243
OY 242 tactatctctcaagcttggcagctcagaatctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 301
DB 244 tactatctctcaagcttggcagctcagaatctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 303
OY 302 gaaatctagaagaatgtggaacaattacaccttctctctgttgggctgtggaatgtacttc 361
DB 304 gaggctcactagaatgtggaacaattacaccttctctctgttgggctgtggaatgtacttc 363
OY 362 aagacagccactctctcgaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 421
DB 364 aagacagccactctctcgaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 423
OY 422 gtagagcgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 481
DB 424 gtagagcgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 483
OY 482 cgggcccctcaggatctctcagctcagctcagctcagctcagctcagctcagctcagctc 541
DB 484 cgggcccctcaggatctctcagctcagctcagctcagctcagctcagctcagctcagctc 543
OY 542 accagcatcctcagctcagctcagctcagctcagctcagctcagctcagctcagctcag 601
DB 544 accagcatcctcagctcagctcagctcagctcagctcagctcagctcagctcagctcag 603
OY 602 gccacatcagctcagctcagctcagctcagctcagctcagctcagctcagctcagctcag 661
DB 604 gccacatcagctcagctcagctcagctcagctcagctcagctcagctcagctcagctcag 663
OY 662 ttcctctctatatacttcccaatgacctcagctcagctcagctcagctcagctcagctcag 721
DB 664 ttcctctctatatacttcccaatgacctcagctcagctcagctcagctcagctcagctcag 723
OY 722 aggtcgaagaagaatgaaatcccttgaagcgaacaagaatgtgtgtgtgtgtgtgtgtgt 781
DB 724 aggtcgaagaagaatgaaatcccttgaagcgaagaagaatgtgtgtgtgtgtgtgtgtgt 783
OY 782 tctagaagaatcagctcagaagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 841
DB 784 tctagaagaatcagctcagaagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 843
OY 842 accccctcctcagctgtgacccggtctctctcagcttctgtgagaagtgagcagaagctcc 901
DB 844 gccccgtctcacaatgacacgcaactctctctcagcttctgtgagaagtgagcagaagctcc 903
```

```

OY 902 gctgctgtgttcaaccttcaaccttcaaccttcaaccttcaaccttcaaccttcaaccttca 961
DB 904 gctgctgtgttcaaccttcaaccttcaaccttcaaccttcaaccttcaaccttcaaccttca 963
OY 962 gtcaaccccatattctataactctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1021
DB 964 gtcaaccccatattctataactctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1023
OY 1022 gtccctcctacgtgcaagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1081
DB 1024 atctctcttcccaacaagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1083
OY 1082 aagatcatcttctgacagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1141
DB 1084 cggaaatcttctctgacagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1143
OY 1142 ttcctctgtcagtcacatcctcacaacacacacacacacacacacacacacacacacac 1201
DB 1144 ttcctctgtcagtcacatcctcacaacacacacacacacacacacacacacacacacac 1203
OY 1202 gtagcactaa 1211
DB 1204 atgtcaagaa 1213
```

RESULT 11
AAAA6022
ID AAAA6022 standard; cDNA; 1248 BP.

XX AAAA6022;
XX 22-AUG-2000 (first entry)

DE Human G protein coupled receptor hRup6 encoding cDNA SEQ ID NO:11.

KW Human; G protein coupled receptor; GPCR; transmembrane receptor;
KW Identification; agonist; screening; therapeutic; pharmaceutical;
KW mutant; ss.

OS Homo sapiens.

PN WO200022131-A2.

PD 20-APR-2000.

XX 13-OCT-1999; 99WO-US24065.

XX 13-OCT-1998; 98US-0170496.

XX 12-NOV-1998; 98US-0108029.

XX 20-NOV-1998; 98US-0109213.

XX 27-NOV-1998; 98US-0110060.

XX 26-FEB-1999; 99US-0120416.

XX 12-MAR-1999; 99US-0123944.

XX 12-MAR-1999; 99US-0123945.

XX 12-MAR-1999; 99US-0123946.

XX 12-MAR-1999; 99US-0123948.

XX 12-MAR-1999; 99US-0123949.

XX 12-MAR-1999; 99US-0123951.

XX 28-MAY-1999; 99US-0136436.

XX 28-MAY-1999; 99US-0136437.

XX 28-MAY-1999; 99US-0136439.

XX 28-MAY-1999; 99US-0137127.

XX 28-MAY-1999; 99US-0137131.

XX 30-JUN-1999; 99US-0141448.

XX 27-AUG-1999; 99US-0151114.

XX 03-SEP-1999; 99US-0152524.

XX 29-SEP-1999; 99US-0156633.

XX 29-SEP-1999; 99US-0156634.

XX 29-SEP-1999; 99US-0156634.

XX (AREN-) ARENA PHARM INC.

XX Behan DP, Lehmann-Brunnsma K, Chalmers DT, Chen R, Dang HT;
PI Gore M, Llaw CM, Lin I, Lowitz K, White C;
XX MPI: 2000-317986/27.
XX P-PSDB: AAB02830.
DR
XX Non-endogenous, human G protein-coupled receptors for screening
PT receptor, inverse or partial agonists useful as therapeutic agents -
XX
XX Example 1: Page 85-86; 187pp; English.
XX
XX The present invention describes transmembrane receptors, preferably
CC human G protein coupled receptors (GPCR), for which the endogenous
CC ligand is unknown (orphan GPCR receptors). More specifically the present
CC invention relates to non-endogenous, constitutively activated versions
CC of a human GPCR. These non-endogenous human GPCRs can be useful for
CC the direct identification of candidate compounds as receptors agonists,
CC inverse agonists or partial agonists for use as pharmaceutical agents.
CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used
CC in the exemplification of the present invention.
XX
S0 Sequence 1248 BP; 269 A; 393 C; 269 G; 317 T; 0 other;

Query Match	64.2%	Score 790.8	DB 21	Length 1248
Best Local Similarly	79.2%	Pred. No. 1.3e-224		
Matches 958	Conservative	0	Mismatches 237	Indels 15
				Gaps 1

Oy	17	tcgaaatcgggaaataatcgaaatagctcttcctcgagatccac-----gattcct	61
Db	4	tcagagatcgggaaaaaacttcgaaatgctcttcctcgatctcaacagcagaatactagaatcca	63
Oy	62	ctcabcgaagtaactcgtaaacagcacagagagtaactctggccaccctcgtctgaaaccacaagcgc	121
Db	64	ctccagaaacaactcgaaacagcacagcagagtaattcgtcttcctctctcgcgaacttcgcgc	123
Oy	122	agtcgaccatacttccttcgcgcgtctcgtctgagcctatgcgtctgattcttcctcgctgggggtcaag	181
Db	124	agccactctcttcctcccccgcgtctcgtctgtgtatctgtacaaattctctgtgtgggggtcaatt	183
Oy	182	ggcaactctctcgtctcgtcatgtatctgtctccgagacatacagaactcttgaaagacaaccacaac	241
Db	184	ggcaaatcgtccgcgtgctcgcgtctcgtatctctcgaagacaacagcctatgaaagagccaccacaac	243
Oy	242	tactatctctcagctctgagcagctccacagatctgctctgctcgtcctctctgaggatctcgtcgt	301
Db	244	tactctctcttcacagctcgtcgcgtctctctcgtacotctcgtcttcctcgtcctctctgaaatctgcctcgt	303
Oy	302	gaaatactcagagatcgtgagacaaatatacctcttcctgctgttgagcgtctggagatgtaacttc	361
Db	304	gagatctatcgtgaagatcgtgcgaaatacctacccctctctctgttcgggcccgtggcgtcctaactc	363
Oy	362	aagacagccccctctccagagacatctgtctctctgacccaattctcagctgacacaagcgtatagc	421
Db	364	aagacagccccctctctgagacccgtgtgtcttcgcctctcaactctccagacatacacacacgtccagc	423
Oy	422	gttagagcgtctatctgtgcacatctgtccaccctcttcgagacaaagctgtgagagcacgcggcgga	481
Db	424	gttggtgcgtctcgtgcacatctctcaaccctcttcctcgcccaaaactgtgcagagcacaccgcgcg	483
Oy	482	cgggcgccctcagaaatccacagcctatgctctgtgagactctctctgtgtgctcttcttcctggccaat	541
Db	484	cgggcgccctcagaaatccctccgcgacatctgcctcgggactcttcctcgcgtcctctctccctgcgccaac	543
Oy	542	accagacatccaatggcgcatactaaagtctccagacactcttcccaacgggtctccctcgtacactgtctca	601
Db	544	accagacatccaatggcgcatactaaagtctccacatactctcccaatgggtctccctcgtccaggtctcg	603
Oy	602	ggccacctcgcacaagtcacacaaaccacatgtgtgtatataactgtatcatccaagctacacagc	661
Db	604	ggccacctcgtatgcatactaaagccacatgtgtgtatctataactatcatccaatcgcagctcattcttc	663

OY	662	ttcctcttctatcatccctcccaatgacccctatagaagtcctctactactatctatgagctc	721
Db	664	ttccatctctctacccctccctcccaatgacctgctatcaatgctctctactactactatgacatc	723
OY	722	agagctgaagagaatgatatcccttgagagcgaaacaagtgbcgtgtagatatctcaagacc	781
Db	724	agactaaagaagaagcaaatctcttgagggcagaatgaaagggaaatgcaaatcttcaagacc	783
OY	782	tcctagaagatcagtcacccaagaatgctgtttgtcttgctcccggtgtttgcacatctcgtg	841
Db	784	tgcagaanaatcagtcacccaagaatgctgtttgtcttgctccagtgctatcgtctg	843
OY	842	accocctccatctgagcgcgctctcttcaagctttgtgaaagatgagcaagctcctg	901
Db	844	gocccgcttccacatctgacccgaactctcttcaagctttgtgaaagatgagcaagctcctg	903
OY	902	gctgctgtgttcaaccctcatcatatgtgatactgaagtgctctcttctatcttgagctcccg	961
Db	904	gctgctgtgttcaaccctcgtccatgtggtgtcgaagtgctctcttctactactgagctccagct	963
OY	962	gtcaaacccattatataaaacctctgtcttgagcgctcttcgggagcgagcctttcgaaatgtc	1021
Db	964	gtcaaacccattatataaaacctctgtcttgagcgcttccagcgagcatcttcgaatgtc	1023
OY	1022	gtcttcocctactctgcaaatgtgtccatctcccgagcatcggcgcaagaaggaactccagccag	1081
Db	1024	atctcttcttcttccacaacaagtggtccatctcccaagtcacagccaaagtcttgcaacctgtccag	1083
OY	1082	aagatcatctctctgacaagaatgacacctgtgagatgacagaagatgcagagccccag	1141
Db	1084	cggaaacatcttcccgacaagaatgtccactgtgtgagctgacacgaagaatctagttcccaa	1143
OY	1142	ttccctgtgctacatccatccacaacaaccaaaccttaccacgagcccccgtgtgcagagag	1201
Db	1144	ttcccatgtcagatccatccatgacaacaactccacctcccaacagccctctctagtgaaacag	1203
OY	1202	gtacacataa	1211
Db	1204	atgtcaagaa	1213

XX	RESULT 12
XX	AAH49526
ID	AAH49526 standard; DNA; 1248 BP.
XX	
AC	AAH49526;
XX	
DT	21-SEP-2001 (first entry)
XX	
DE	Human GTP-binding protein-coupled receptor GPRV39 coding sequence.
XX	
KM	GTP-binding protein-coupled receptor; neuroprotective; immunomodulatory;
KM	muscular; urinary; circulatory; anorectic; gene therapy; human;
KW	guanosine triphosphate; G-protein; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200148189-A1.
XX	
PD	05-JUL-2001.
XX	
PF	28-DEC-2000; 2000WO-JP09409.
XX	
PR	28-DEC-1999; 99JP-0375152.
PR	31-MAR-2000; 2000JP-0101339.
PR	23-MAY-2000; 2000JP-0155978.
XX	
PA	(HELI-) HELIX RES INST.
XX	
PI	Matsumoto S, Oda T, Saito Y, Morikawa N, Yoshida K, Suwa M;
PI	Sugiyama T;
XX	
DR	WPI; 2001-425663/45.

P-PSDB: AAC64297.

Family of guanosine triphosphate binding protein coupled receptors and genes encoding them for treatment and prevention of diseases associated with these receptors.

Claim 1: Page 105, 137pp; Japanese.

The present sequence is the coding sequence for a human guanosine triphosphate (GTP)-binding protein coupled receptor. The receptor is useful for the investigation, diagnosis, treatment and prevention of diseases associated with GTP-binding protein-coupled receptors, including neurological, circulatory, digestive system, immune system, muscle and urinary system disorders. GTP-binding proteins are also known as G-proteins.

Sequence 1248 BP; 269 A; 393 C; 269 G; 317 T; 0 other:

Query Match 64.2%; Score 790.8; DB 22; Length 1248;
Best Local Similarity 79.2%; Pred No. 1.3e-224;
Matches 958; Conservative 0; Mismatches 237; Indels 15; Gaps 1;

```

Oy 17 tcagtaatgggaaacttgaaaaatgcttcctgataccac-----gactct 61
Db 4 tcaggatggaanaaacttcgaatgctctcctgatactacagcagaactagaagatcca 63
Oy 62 ctatgaatgaacttgaaacagacagagagatcttgccacctggttgagcccaagcgc 121
Db 64 ttccagaagaacacacagacagcagagagatctgctcctccctcctgctgctgagcgc 123
Oy 122 agtgaacctacccttcgagctgctgagcctatgagcctgactcctccgagggggaatg 181
Db 124 agccactctctctcccccggctgctggtatgtagtgcaatttcttggtggggatcat 183
Oy 182 ggcaatctctctggtgctgctgctgctgctgctgctgctgctgctgctgctgctgct 241
Db 184 ggcatactctctggtgctgctgctgctgctgctgctgctgctgctgctgctgctgct 243
Oy 242 tactatctctctctgagctgagctgagctgagctgagctgagctgagctgagctgag 301
Db 244 tactatctctctctgagctgagctgagctgagctgagctgagctgagctgagctgag 303
Oy 302 gaatactacagagaatgctgacacaaatccctctctctgctgagcctgagctgagct 361
Db 304 gaagctctatgagatgctgagcgaacacacacacacacacacacacacacacacac 363
Oy 362 aagacagccctctctgagacgtgctgctgctgctgctgctgctgctgctgctgctg 421
Db 364 aagacagccctctctgagacgtgctgctgctgctgctgctgctgctgctgctgctg 423
Oy 422 gtagagcagctatgctgagcctatgctgagcctctctgagcagctgagcagcagc 481
Db 424 gtagagcagcctatgctgagcctatgctgagcctctctgagcagcctgagcagc 483
Oy 482 cgggcccctacagagatccctcagcctagctgagcctctctgagcctctctgagc 541
Db 484 cgggcccctacagagatccctcagcctagctgagcctctctgagcctctctgagc 543
Oy 542 accagacatccatgagcctacagcctacagcctacagcctacagcctacagcctac 601
Db 544 accagacatccatgagcctacagcctacagcctacagcctacagcctacagcctac 603
Oy 602 ggcacacagcagctgagcctacagcctacagcctacagcctacagcctacagcctac 661
Db 604 ggcacacagcagctgagcctacagcctacagcctacagcctacagcctacagcctac 663
Oy 662 ttccctctctacacacacacacacacacacacacacacacacacacacacacacac 721
Db 664 ttccctctctacacacacacacacacacacacacacacacacacacacacacacac 723
Oy 722 aggcctgaagagagatgaatccctctgagcagacaaagctgagctgtaatactacag 781

```

```

Db 724 agactaaagaagacaaatctcttgagcgagatgaaaggaatgcaaatatcaagaacc 783
Oy 782 tctagaagaatcagctacacagagatgctgttctgttctgttctgttctgttctgtt 841
Db 784 tgcagaanaatcagctacacagagatgctgttctgttctgttctgttctgttctgtt 843
Oy 842 accccctcagctgagcagcagcctctctcagcctctctcagcctctctcagcctct 901
Db 844 gccccgttccacatgacacagcagcctctctcagcctctctcagcctctctcagcct 903
Oy 902 gctgctgtgttccacatgacacagcagcctctctcagcctctctcagcctctct 961
Db 904 gctgctgtgttccacatgacacagcagcctctctcagcctctctcagcctctct 963
Oy 962 gtcaacccatatactatacctcctgctctcagcctctctcagcctctctcagcct 1021
Db 964 gtcaacccatatactatacctcctcagcctctctcagcctctctcagcctctcag 1023
Oy 1022 gtctccctacacagcagcagcctctcagcctctctcagcctctctcagcctcag 1081
Db 1024 atctctctcttccacacacagcagcctctcagcctctctcagcctctcagcctcag 1083
Oy 1082 aagatcatctcttgaacagatgctacacacagcagcctctcagcctctcagcctcag 1141
Db 1084 cggacatctctctgacagatgctacacacagcagcctctcagcctctcagcctcag 1143
Oy 1142 ttccctgtgtcagcagcagcctctcagcctctctcagcctctctcagcctctcag 1201
Db 1144 ttccctgtgtcagcagcagcctctcagcctctctcagcctctctcagcctctcag 1203
Oy 1202 gtacataaa 1211
Db 1204 atgtcaagaa 1213

```

```

RESULT 13
AAH43075
ID AAH43075 standard; DNA; 1245 BP.
XX
AC AAH43075;
XX
DT 15-OCT-2001 (first entry)
XX
DE Nucleotide sequence of a human TGR-1 protein.
XX
KW TGR-1; neuromedin U; hypertension; stress disease; ss.
XX
OS Homo sapiens.
XX
FH Key 1.1245 Location/Qualifiers
FT CDS /tag= a "TGR-1"
FT FT /product= "TGR-1"
FT FT /note= "no termination codon given"
PN WO200157524-A1.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-IP00746.
XX
PR 04-FEB-2000; 2000JP-0032773.
PR 24-FEB-2000; 2000JP-0052252.
PR 30-MAR-2000; 2000JP-0097896.
PR 19-JUN-2000; 2000JP-0187536.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PI Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;
PI Okubo S;
XX
DR WPI: 2001-488917/53.
DR P-PSDB: AAC63366.

```

XX Identifying predicted or actual structures of two or more members of a
 PT chemical or physical library by mass spectrometry comprising
 PT correlating molecular mass measurements of two or more members with a
 PT shared chemical history -
 XX
 PS Disclosure: Page 85-87; 95pp; Japanese.
 XX
 CC The present sequence encodes a human TGR-1 protein. The specification
 CC describes a method of screening a compound, which is capable of binding
 CC properties of neuromedin U to TGR-1. The method is useful for screening
 CC preventatives and remedies for hypertension, stress diseases, etc..
 CC TGR-1 antagonists are also useful for treating the same diseases.
 CC
 XX Sequence 1245 BP: 268 A; 394 C; 268 G; 315 T; 0 other;

Query Match 64.0%; Score 787.6; DB 22; Length 1245;
 Best Local Similarity 79.0%; Pred. NO. 1.le-223;
 Matches 956; Conservative 0; Mismatches 239; Indels 15; Gaps 1;

Oy 17 tcaagtaatgggaaactggaaatgtcttcctgagtcac-----gactct 61
 Db 4 tcaaggaatggaataacttcagaaatgtcttcctgagtcacacagcagaatactagaatcca 63
 Oy 62 ctcaatgaatctctgaaacagacagagagatctctgcccactgtgtgaccacagcgc 121
 Db 64 ttcacgaataacacactgaaacagacagagagatctctgcccactgtgtgaccacagcgc 123
 Oy 122 agtgaactacccctcgcgtgtctgtgagcctatgacgtcatctcttcgtgagggttaag 181
 Db 124 agccactctctctccctcgcgtgtctgtgagcctatgacgtcatctcttcgtgagggttaag 183
 Oy 182 ggcaactctctctggtgagtgatggtgagtcgcagacacagacttgaacacacccacac 241
 Db 184 ggcaatgctcgggtgtgagtcggtgagtcgacacagcagcagatgaagacgcccacac 243
 Oy 242 tactatctctcagcttgagcagctcagacatctgctgctcctcttgagggtgctctg 301
 Db 244 tactatctctcagcttgagcagctcagacatctgctgctcctcttgagggtgctctg 303
 Oy 302 gaaatctcagagatgtgacacaaatcactctctctgtctcgagcgtgtgaggtactctc 361
 Db 304 gaggctcactgagatgtgagcgaactacactctctctgtctcgagcgtgtgaggtactctc 363
 Oy 362 aagacagccctctcagaaactgtgtgcttgcctcactctcagtgatcacaacaggttgc 421
 Db 364 aagacagccctctcagaaactgtgtgcttgcctcactctcagtgatcacaacaggttgc 423
 Oy 422 gtaagagcctatgtgacatctgacaccccttcgcagacacagctgagagcagcgagcga 481
 Db 424 gtagagcgtctcgtgacatctgacaccccttcgcagacacagctgagagcagcgagcga 483
 Oy 482 cgggacctcagatctcagcctcagctcgtgagctctcgtgtgctcttcttcttgcacat 541
 Db 484 cgggacctcagatctcagcctcagctcgtgagctctcgtgtgctcttcttcttgcacac 543
 Oy 542 accgagcactcagtgatcagaatgtcagacacttctccacagcggtctcctgactgcctca 601
 Db 544 accgagcactcagtgatcagaatgtcagacacttctccacagcggtctcctgactgcctc 603
 Oy 602 gccacctcagacatcacaac 661
 Db 604 gccacctcagacatcacaac 663
 Oy 662 ttcctctctacatctcctccac 721
 Db 664 ttcctctctacatctcctccac 723
 Oy 722 aggcctgaagagagatgaatctcctctgagcgaaacaaagtgcgtgtgaatctcagacac 781
 Db 724 agactaaagaaagacaaatctcctctgagcgagatgaggaatgcaaatatctcagacac 783

Oy 782 tctagaagaatcagtcacacacagatgctgtgtctgtgtctgtgtgtgtgtgtgtgtgtgt 841
 Db 784 tgcagaataatcagtcacacacagatgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 843
 Oy 842 accccctcactgtagcagcgtctcttcacagcttctgtgaaagatgacagatctcctg 901
 Db 844 gccccctcactgtagcagcgtctcttcacagcttctgtgaaagatgacagatctcctg 903
 Oy 902 gctgtgtgttcacatctcactatcagtgatgagtgatgagtgatgagtgatgagtgatgag 961
 Db 904 gctgtgtgttcacatctcactatcagtgatgagtgatgagtgatgagtgatgagtgatgag 963
 Oy 962 gtaacacacatctcactatcagtgatgagtgatgagtgatgagtgatgagtgatgagtgat 1021
 Db 964 gtaacacacatctcactatcagtgatgagtgatgagtgatgagtgatgagtgatgagtgat 1023
 Oy 1022 gctcctcactcagacatgtagtgacatcccgacacacacacacacacacacacacacac 1081
 Db 1024 atctctcttcac 1083
 Oy 1082 aagatcactctctgacagaaatgtagctgtgtgagcgtgacagagagatgagcgcacac 1141
 Db 1084 cggaaacatctctcctgacagaaatgtagctgtgtgagcgtgacagagagatgagcgcacac 1143
 Oy 1142 ttcctctgtagacatctcactatcacaacacacacacacacacacacacacacacacac 1201
 Db 1144 ttcctctgtagacatctcactatcacaacacacacacacacacacacacacacacacac 1203
 Oy 1202 gtacacataa 1211
 Db 1204 atgtcaagaa 1213

RESULT 14

AAE80326 standard; DNA; 1594 BP.

AAE80326;

29-JUN-2001 (first entry)

Splice variant of G-protein coupled receptor IGS4A long version cDNA.

Human: G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
 nervous system disorder; psychiatric disorder; parkinson's disease;
 episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
 cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
 motility disorder; myocardial infarction; hypertension; dyslipidemia;
 gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
 inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
 gynecological disorder; ss.

Homo sapiens.

Location/Qualifiers

55..945

/tag= a

/product=

IGS4A long version"

MO200125269-A2.

12-APR-2001.

25-SEP-2000; 2000WO-EP09584.

24-SEP-1999; 99EP-0203140.

24-SEP-1999; 99NL-1013140.

28-JUL-2000; 2000EP-0202683.

31-JUL-2000; 2000US-0222047.

(SOLV) SOLVAY PHARM BV.

PI Deleersnyder W, Berger C, Loeken C, Nys G, Venema J:
 XX WPI: 2001-273568/28.
 DR P-PSDB: AAB67806.
 XX
 PT New G-protein coupled receptors and the polynucleotides encoding them,
 PT useful for preventing, ameliorating or correcting nervous system
 PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
 PT or cancers
 XX
 PS Example 1b: Page 90-92; 102pp; English.
 XX
 CC The present sequence encodes a splice variant of the long version of a
 CC human G-protein coupled receptor designated IGS4A. IGS4 exists in two
 CC polymorphic forms, IGS4A and IGS4B. The IGS4 receptors and IGS4
 CC polynucleotides are useful for preventing, ameliorating or correcting
 CC dysfunctions or diseases. These diseases include peripheral nervous
 CC system, psychiatric and central nervous system disorders
 CC (e.g. schizophrenia, episodic paroxysmal anxiety, phobia, Parkinson's
 CC disease, migraine, epilepsy, bulimia or stroke), cardiovascular diseases
 CC (e.g. heart failure, angina pectoris, myocardial infarction or
 CC hypertension), dyslipidemias, obesity, emesis, gastrointestinal
 CC disorders (e.g. inflammatory bowel disease or motility disorders),
 CC osteoporosis, inflammations, infections (e.g. bacterial, fungal,
 CC protozoan or viral), pain, cancers, immune disorders, allergies,
 CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
 CC effective with regard to disorders of the nervous system, including the
 CC central and peripheral nervous systems, disorders of the gastrointestinal
 CC system, cardiovascular system, skeletal muscle, thyroid, lung or
 CC genitourinary system, or immunological disease. The IGS4 polynucleotides
 CC are useful as diagnostic reagents for detecting under-expression,
 CC overexpression or altered expression of IGS4.
 CC
 XX Sequence 1594 BP: 370 A: 473 C: 332 G: 419 T: 0 other:

Query Match 55.0%; Score 677.6; DB 22: Length 1594;
 Best Local Similarity 74.8%; Pred. No. 66-191;
 Matches 914; Conservative 0; Mismatches 229; Indels 79; Gaps 2;

QY 5 tggatttaagctcgttaatggaacttgaaatgcttcctcggtacc----- 55
 DB 46 tggattttaaagctcgttaatggaacttgaaatgcttcctcggtaccagcagaa 105
 QY 56 -----gacccctcagctgaactgaacagcagagagatgctgcacacgtgt 109
 DB 106 ctgaagatcattccagaaaccccgaaacagcagagagatgctgcacacgtgt 165
 QY 110 ggaaccaagcagtgacatccctccggtgtctgtgacctagcgtatctctc 169
 DB 166 ggaacctcgagcagcactctctcccggtgtctgtgacctagcgtatctctc 225
 QY 170 gtggggtaatggcactctctggtgtcagtgtgattgtccgacatagaatttgaag 229
 DB 226 gtgggggtatctggaatgctcgtgtgtcgtgtgattgtccgacatagaatttgaag 285
 QY 230 aacaccacaactactctctcagctgagctgcagatctgctgtgctcgtctc 289
 DB 286 acgcccacaactactactctctcagctgagctgtcgtcgtcgtcgtcgtcctc 345
 QY 290 ggaatgctcctggaatactcagagatgtgacaaatcactcctctgttgggacctgtg 349
 DB 346 ggaatgctcctggaatgctcagatgtgtgacaaatcactcctctgttgggacctgtg 405
 QY 350 ggaatgctcctggaatgctcagatgtgtgacaaatcactcctctgttgggacctgtg 409
 DB 406 ggaatgctcctggaatgctcagatgtgtgacaaatcactcctctgttgggacctgtg 465
 QY 410 acaaggttagcgttagagcgtatgtgacatgttcacaccttcggaagcagggag 469
 DB 466 acaaggttagcgttagagcgtatgtgacatgttcacaccttcggaagcagggag 525
 QY 470 agcaacgagcagcagcctcagatcctcagcctagctgtgagcttctctgtgtctt 529

DB 526 agcacccgagcgcggccctcagatcctcggcagctgcgtcctcctc 585
 QY 530 tcttggcccaatacagatcctcagatcctcggcagctgcgtcctcctc 589
 DB 586 tccctggcccaatacagatcctcagatcctcggcagctgcgtcctcctg 645
 QY 590 gtaacctgctcagcagcctcagatcctcggcagctgcgtcctcctc 649
 DB 646 gtaacctgctcagcagcctcagatcctcggcagctgcgtcctcctc 705
 QY 650 caagctaccagctcctcctcagatcctcggcagctgcgtcctcctc 709
 DB 706 caagctaccagctcctcctcagatcctcggcagctgcgtcctcctc 765
 QY 710 cctcagcagcctcagatcctcggcagctgcgtcctcctcctc 769
 DB 766 cctcagcagcctcagatcctcggcagctgcgtcctcctcctc 825
 QY 770 attcacaagcctcagatcctcggcagctgcgtcctcctcctc 829
 DB 826 attcacaagcctcagatcctcggcagctgcgtcctcctcctc 865
 QY 830 gccatctgctgaacccctcagatcctcggcagctgcgtcctcctc 889
 DB 866 -----ccttggagagatg 881
 QY 890 acaagctcctgctcctcctcagatcctcggcagctgcgtcctcctc 949
 DB 882 agtgaatcctgctcctcctcagatcctcggcagctgcgtcctcctc 941
 QY 950 ctgaatcctgctcctcctcagatcctcggcagctgcgtcctcctc 1009
 DB 942 ctgaatcctgctcctcctcagatcctcggcagctgcgtcctcctc 1001
 QY 1010 ttcgaatgctcctcctcagatcctcggcagctgcgtcctcctc 1069
 DB 1002 ttcgaatgctcctcctcagatcctcggcagctgcgtcctcctc 1061
 QY 1070 cctcagcagcctcagatcctcggcagctgcgtcctcctcctc 1129
 DB 1062 caactgctcagcagcctcctcagatcctcggcagctgcgtcctcctc 1121
 QY 1130 gcaagcctcagcagcctcctcagatcctcggcagctgcgtcctcctc 1189
 DB 1122 atagtgcccaatcctcagatcctcggcagctgcgtcctcctc 1181
 QY 1190 tgtgacagagatcactataa 1211
 DB 1182 tctagtgacagatgtaagaa 1203

RESULT 15
 AAF80327
 ID AAF80327 standard; DNA: 1594 BP.
 XX
 AC AAF80327;
 XX
 DT 29-JUN-2001 (first entry)
 XX
 DE Splice variant of G-protein coupled receptor IGS4A short version cDNA.
 DE
 XX Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
 XX nervous system disorder; psychiatric disorder; Parkinson's disease;
 XX episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
 XX cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
 XX motility disorder; myocardial infarction; hypertension; dyslipidemia;
 XX gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
 XX inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
 XX gynecological disorder; ss.
 XX
 OS Homo sapiens.
 XX

Job time: 6395 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:19:18 ; Search time 43.26 Seconds
(without alignments)
1335.589 Million cell updates/sec

Title: US-09-609-146-25
2076
Perfect score: 1 MGLKLNASWTHDPLMKRYLNS.....GOSSIHNTLTAPCAGEVP 395
Sequence:

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: SPTRMBL_17.*
 - 2: sp_archaea.*
 - 3: sp_bacteria.*
 - 4: sp_fungi.*
 - 5: sp_human.*
 - 6: sp_invertebrate.*
 - 7: sp_mammal.*
 - 8: sp_mhc.*
 - 9: sp_organelle.*
 - 10: sp_phage.*
 - 11: sp_plant.*
 - 12: sp_protent.*
 - 13: sp_virus.*
 - 14: sp_vertebrate.*
 - 15: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2076	100.0	395	11	Q9JIB1
2	2063	99.4	395	11	Q9ESQ4
3	1650.5	79.5	412	4	Q9NRAB
4	1650.5	79.5	415	4	Q9GZ04
5	952.5	45.9	403	4	Q43664
6	952.5	45.9	426	4	Q9HB89
7	921.5	44.4	402	11	Q9JIB2
8	916.5	44.1	412	11	Q9JIB2
9	840	40.5	405	11	Q55040
10	564.5	27.2	428	5	Q9VFNA
11	546.5	26.3	595	5	Q9VFNA
12	525.5	25.3	660	5	Q9VFNA
13	495.5	23.9	418	5	Q9VFNA
14	476.5	22.0	363	13	Q93413
15	470	22.6	374	13	Q93412
16	457.5	22.0	378	5	Q18701
17	401	19.3	401	13	Q9DDR0
18	400.5	19.3	454	4	Q9H573
19	399.5	19.2	400	6	Q9MYW9

20	398	19.2	397	13	Q9DDR1	Q9ddr1 xenopus lae
21	394	19.0	398	4	Q9UDE7	Q9ude7 homo sapien
22	390	18.8	504	5	Q9VX8	Q9vxx8 drosophila
23	388.5	18.7	398	4	Q9UDE6	Q9ude6 homo sapien
24	388	18.7	404	13	Q9DFB0	Q9dfb0 catostomus
25	387	18.6	393	11	Q9R1M0	Q9rlm0 mus musculus
26	387	18.6	401	11	Q9R1L9	Q9rl19 mus musculus
27	387	18.6	438	11	Q9R0D1	Q9r0d1 mus musculus
28	387	18.6	444	11	Q9J1Y1	Q9j1y1 mus musculus
29	384.5	18.5	452	11	Q9JKN0	Q9jkn0 mus musculus
30	381.5	18.4	394	5	Q9U721	Q9u721 drosophila
31	379	18.3	377	13	Q9BU14	Q9bu14 brachydanio
32	376.5	18.1	383	13	Q42324	Q42324 catostomus
33	371.5	17.9	513	13	Q9DFA9	Q9dfa9 catostomus
34	368.5	17.8	367	13	Q9PVC0	Q9pvc0 carassius a
35	368.5	17.8	367	13	Q9PVC9	Q9pvc9 carassius a
36	366	17.6	384	13	Q9BUH1	Q9buh1 brachydanio
37	365	17.6	559	13	Q9JK40	Q9jkl40 mus musculus
38	360.5	17.4	385	11	Q9JKA0	Q9jka0 mus musculus
39	360	17.3	388	4	Q9J1Y1	Q9j1y1 homo sapien
40	354.5	17.1	352	11	Q9B820	Q9b820 rattus norv
41	354.5	17.1	352	11	Q9R297	Q9r297 rattus norv
42	354.5	17.1	352	11	Q9QW3	Q9qw3 rattus norv
43	354	17.1	425	6	Q9M2U2	Q9m2u2 oryctolagus
44	354	17.1	429	6	Q9M2U3	Q9m2u3 oryctolagus
45	347.5	16.7	466	11	Q9WU25	Q9wu25 cavia porce

ALIGNMENTS

RESULT 1

Q9JIB1 PRELIMINARY; PRT; 395 AA.

AC Q9JIB1;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE NEUROMEDIN U RECEPTOR 2.

GN NMUR2K.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY;

RX MEDLINE=20351041; PubMed=10894543;

RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M., Zeng Z., Williams D.L., Feigmet S.D., Nunes C.N., Murphy B., Hrenuk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austin C.P., Caskey T., van der Ploeg L.H.T., Liu Q.;

RT "Identification of receptors for neuromedin U and its role in feeding."

RT Nature 406:70-74(2000).

RL EMBL: AF242875; AAF82756.1; -

DR InterPro: IPR000276; GPCR_Rhodopsn.

DR Pfam: PF00001; 7tm_1.1.

DR PRINTS: PR00237; GPCR_Rhodopsn.

DR PROSITE: PS00237; G_PROTEIN_REC_F1_1; UNKNOWN_1.

DR PROSITE: PS50262; G_PROTEIN_REC_F1_2; 1.

KW Receptor.

SO SEQUENCE 395 AA; 44722 MW; 01D3765B5D5355C0 CRC64;

Query Match 100.0%; Score 2076; DB 11; Length 395;
Best Local Similarity 100.0%; Pred. No. 1.3e-151;
Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLKLNASWTHDPLMKRYLNSSTEYLAHLGCPKRSDSLSPVSAVALIFLVGWNLLVCM 60
Db 1 MGLKLNASWTHDPLMKRYLNSSTEYLAHLGCPKRSDSLSPVSAVALIFLVGWNLLVCM 60

```

QY 61 VIVRHOTLKTPTNYLFSFLAVSDLLVLLGMPLEIYEMHNPFLPGVCYFKTALFET 120
    |||||||
DB 61 VIVRHOTLKTPTNYLFSFLAVSDLLVLLGMPLEIYEMHNPFLPGVCYFKTALFET 120
QY 121 VCFASLSTVTTVSERYVAIVHFPRAKLESTRRRALRILSLVWSFSVFLSPTMSIHGIK 180
    |||||||
DB 121 VCFASLSTVTTVSERYVAIVHFPRAKLESTRRRALRILSLVWSFSVFLSPTMSIHGIK 180
QY 181 FOHPNGSSVPGSATCTVTKPMVYNLIIOATSFLEFIILPMTLISVLYLMLGRKDES 240
    |||||||
DB 181 FOHPNGSSVPGSATCTVTKPMVYNLIIOATSFLEFIILPMTLISVLYLMLGRKDES 240
QY 241 LEANKAVANIHRRPSRKSVTKMLFVLVLFALICWTPFHVDRLFESFVEEMTESLAAVENLI 300
    |||||||
DB 241 LEANKAVANIHRRPSRKSVTKMLFVLVLFALICWTPFHVDRLFESFVEEMTESLAAVENLI 300
QY 301 HVGSGVFYLLSSAVNPITVYLLSRFRRAFRNVSPCKKCHPRHPOGPPAKIIFLTE 360
    |||||||
DB 301 HVGSGVFYLLSSAVNPITVYLLSRFRRAFRNVSPCKKCHPRHPOGPPAKIIFLTE 360
QY 361 CHLVELTEDAGPOPGOSSIHNTNLITAPCAGEVP 395
    |||||||
DB 361 CHLVELTEDAGPOPGOSSIHNTNLITAPCAGEVP 395

```

RESULT 2

```

Q9ES04 PRELIMINARY: PRT: 395 AA.
AC Q9ES04: 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, last annotation update)
DE G PROTEIN-COUPLED RECEPTOR TGR-1.
GN TGR-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RX PubMed=10887190;
RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
RA Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
RA Fujino M.;
RT "Identification and Functional Characterization of a Novel Subtype of
RT Neuregulin U Receptor."
RL J. Biol. Chem. 275:29528-29532(2000).
DR EMBL: AB041229; BAB13722.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHHODOPSN.
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR_FL_1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECEPTOR_FL_2; 1.
KW Receptor.
SQ SEQUENCE 395 AA: 44756 MW: EABD6D36E9C355DA CRC64:

```

Query Match 99.4%; Score 2063; DB 11; Length 395;
 Best Local Similarity 99.2%; Pred. No. 1,3e-150;
 Matches 392; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MGLLENASWTHIDPLMKYLNSTEEYLAHLGCPKRSDDLSPVSAYALIFLVGMGNLLVCM 60
    |||||||
DB 1 MGLLENASWTHIDPLMKYLNSTEEYLAHLGCPKRSDDLSPVSAYALIFLVGMGNLLVCM 60
QY 61 VIVRHOTLKTPTNYLFSFLAVSDLLVLLGMPLEIYEMHNPFLPGVCYFKTALFET 120
    |||||||
DB 61 VIVRHOTLKTPTNYLFSFLAVSDLLVLLGMPLEIYEMHNPFLPGVCYFKTALFET 120
QY 121 VCFASLSTVTTVSERYVAIVHFPRAKLESTRRRALRILSLVWSFSVFLSPTMSIHGIK 180
    |||||||
DB 121 VCFASLSTVTTVSERYVAIVHFPRAKLESTRRRALRILSLVWSFSVFLSPTMSIHGIK 180

```

```

QY 181 FOHPNGSSVPGSATCTVTKPMVYNLIIOATSFLEFIILPMTLISVLYLMLGRKDES 240
    |||||||
DB 181 FOHPNGSSVPGSATCTVTKPMVYNLIIOATSFLEFIILPMTLISVLYLMLGRKDES 240
QY 241 LEANKAVANIHRRPSRKSVTKMLFVLVLFALICWTPFHVDRLFESFVEEMTESLAAVENLI 300
    |||||||
DB 241 LEANKAVANIHRRPSRKSVTKMLFVLVLFALICWTPFHVDRLFESFVEEMTESLAAVENLI 300
QY 301 HVGSGVFYLLSSAVNPITVYLLSRFRRAFRNVSPCKKCHPRHPOGPPAKIIFLTE 360
    |||||||
DB 301 HVGSGVFYLLSSAVNPITVYLLSRFRRAFRNVSPCKKCHPRHPOGPPAKIIFLTE 360
QY 361 CHLVELTEDAGPOPGOSSIHNTNLITAPCAGEVP 395
    |||||||
DB 361 CHLVELTEDAGPOPGOSSIHNTNLITAPCAGEVP 395

```

RESULT 3

```

Q9NRA6 PRELIMINARY: PRT: 412 AA.
AC Q9NRA6: 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, last annotation update)
DE NEURONMEDIN U RECEPTOR 2 (G PROTEIN-COUPLED RECEPTOR TGR-1).
GN NMUR2 OR TGR-1.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=20351041; PubMed=10894543;
RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
RA Zeng Z., Williams D.L., Feigener S.D., Nunes C.N., Murphy B.,
RA Stair J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McKee K.K.,
RA Hreniuk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austin C.P.,
RA Caskey T., van der Ploeg L.H.T., Liu Q.;
RT "Identification of receptors for neuregulin U and its role in
RT feeding."
RL Nature 406:70-74(2000).
RN (2)
RP SEQUENCE FROM N.A.
RX PubMed=10887190;
RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
RA Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
RA Fujino M.;
RT "Identification and Functional Characterization of a Novel Subtype of
RT Neuregulin U Receptor."
RL J. Biol. Chem. 275:29528-29532(2000).
DR EMBL: AF242874; AAF82755.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHHODOPSN.
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR_FL_1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECEPTOR_FL_2; 1.
KW Receptor.
SQ SEQUENCE 412 AA: 47450 MW: 33D6D14A841B63E CRC64:

```

Query Match 79.5%; Score 1650.5; DB 4; Length 412;
 Best Local Similarity 79.4%; Pred. No. 5.8e-119;
 Matches 312; Conservative 36; Mismatches 40; Indels 5; Gaps 1;

```

QY 1 MGLLENASWTHIDPLMKYLNSTEEYLAHLGCPKRSDDLSPVSAYALIFLVGMGNLLVCM 55
    |||||||
DB 1 MGLLENASWTHIDPLMKYLNSTEEYLAHLGCPKRSDDLSPVSAYALIFLVGMGNLLVCM 55
QY 61 VIVRHOTLKTPTNYLFSFLAVSDLLVLLGMPLEIYEMHNPFLPGVCYFKTALFET 120
    |||||||
DB 61 VIVRHOTLKTPTNYLFSFLAVSDLLVLLGMPLEIYEMHNPFLPGVCYFKTALFET 120

```

Qy	116	ALFETVVCASLSTVTVSVVEYVAIVHPFRAKLESTRARALRLLSLWMSVYFSLPNTS	175
Db	121	ALFETVVCASLSTVTVSVVEYVAIVHPFRAKLQSTRARALRLGLVWGSVYFSLPNTS	180
Qy	176	IHGKIFQHPFPGSSVPASACTVYKRPMMVNLIIQATSEFLYILPMTLSVLEYLGLRL	235
Db	181	IHGKIFHFHPGSLVPASACTVYKRPMMIINFIQVTSFLFLYILPMTVLSVLEYLMAALRL	240
Qy	236	KRDSLEANKAVAVIHRPSKRSYTKMLFVLYLFAICTWTFPHYDRLEFFSVYEEMTSLAA	295
Db	241	KKDSLEADEECNNAIQRCSRVSXKMLEVLYLFAICTMAEFLHIDRLEFFSVYEEMSSLLA	300
Qy	296	VFNLIHVYSGVFYLLSSAVNPILINLSRRFRFAIRNVVSPCTCKWCHRRHPQPAOKI	355
Db	301	VFNLIHVYSGVFYLLSSAVNPILINLSRRFOAAFOVIVISFHKQVHNSQIDPRLPQAQKN	360
Qy	356	IFLFTCHLVELTDEBAGPOFPQOSSIHNTNLTTA	388
Db	361	IFLFTCHFVELTDEDIGPOFPQSSMHNHSHLPYA	393

RESULT	4
09GZ04	
ID	09GZ04
AC	09GZ04
DT	01-MAR-2001 (TREMBLrel. 16, Created)
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE	NEUROMEDIN U RECEPTOR 2 (NEUROMEDIN U RECEPTOR- <i>TYPE 2</i>).
GN	NMR2.
OS	Homo sapiens (human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_Taxid=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=20430668; PubMed=108993166;
RA	Raddatz R., Wilson A.E., Artymyshyn R., Bonini J.A., Borowsky B.,
RA	Bojcu L.W., Zhou S., Kouranova E.V., Nagorny R., Guevarra M.S.,
RA	Dal M., Lerman G.S., Vayase P.J., Branchek T.A., Gerald C., Forray C.,
RA	Adham N.;
RT	"Identification and Characterisation of Two Neuremedin U Receptors
RT	Differentially Expressed in Peripheral Tissues and the Central Nervous
RT	System."
RL	J. Biol. Chem. 275:32452-32459(2000).
RN	[2]
RP	SEQUENCE FROM N.A.
RX	PubMed=11010960;
RA	Shen L., Qiao X., Crona J.H., Behan J., Wang S., Laz T., Bayne M.,
RA	Gustafson E.L., Monsma F.J., Jr., Hedrick J.A.;
RT	"Identification of a Novel Neuremedin U Receptor Subtype Expressed In
RT	the Central Nervous System."
RL	J. Biol. Chem. 275:39482-39486(2000).
RN	[3]
RP	SEQUENCE FROM N.A.
RA	Pang L., Wang S., Laz T., Hedrick J.A.;
RL	Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
DR	EMBL; AF272363; AMG24794.1; -
DR	EMBL; AF292402; AAC603064.1; -
DR	InterPro: IPR000716; GPCR_Rhodopsn.
DR	Pfam: PF00001; T1m.1; 1.
DR	PRINTS: PR00237; GPCRHOOPS.
DR	PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR	PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
SW	RECEPT.
FW	SEQUENCE 415 AA; 47725 MW; C2BACD84B313390F CRC64;

Query Match	79.5%;	Score 1650.5;	DB 4;	Length 415;
Best Local Similarity	79.4%;	Pred. No. 5.8e-119;		
Matches 312;	Conservative 36;	Mismatches 40;	Indels 5;	Gaps 1,

Qy	1	MCKLEMASMih-----DPLMKYLMSTBEYLHLCGPKRSDSLPVSVAUYALPLUYVMGN	55
Db	4	MKLIOMASMIYOQKLEDFPQHLNSTEEXYLAFLCGPRSHFELPVSUUYVPIFWGVIGN	63
Qy	56	LLVCWAVIRHOGTLPKPTUYUFLFSLAVSDLLVLLGMPLEIEMHNNPFLGPGVCYFKT	115
Db	64	VLYCVLVIIOHQAMKPTPTUYFLFSLAVSDLLVLLGMPLEIEMHNNPFLGPGVCYFKT	123
Qy	116	ALFETVCASILSVTTVSVERVYALVHDFRAKLESTRRALRLISLVMSFSVVSFLPNTS	179
Db	124	ALFETVCASILSVTTVSVERVYALVHDFRAKLOSTRRALRLIGIWMGFSVVSFLPNTS	183
Qy	176	IHGIFQHPFPNPNSSVPGSATCTYVKPMVYUMLLIQATSEFLYILPMLISVLYLMLRL	235
Db	184	IHGIFHFPFNPNSSVPGSATCTYVKPMVYUMLLIQATSEFLYILPMLISVLYLMLRL	243
Qy	236	KRDESELEANKVAVNIHRSBRKSVTKMLPVLVLVAICGTPRHVDRLFFSFVEEWBTESTLA	295
Db	244	KKDKSELEADEGANIQRCRCRSYVNMFLVLYLVLAICAPHRIDLRFSEEWBSESLAA	303
Qy	296	VENLIHVYSGVEFYLSAANPILYNLSRRPRAFRUNVSPDKCMCHRNHPQGRPAOKI	355
Db	304	VENLIHVYSGVEFYLSAANPILYNLSRRQAFQNYVSISSFKHQMWSHQDHPQLPRAQRN	363
Qy	356	IFLTCHELVLELTDAGPOFPQOSSIHNTNLTTA	388
Db	364	IFLTCHELVLELTDIGPOFPQOSSMHNNSHPTA	396

RESULT	5			
ID	043664	PRELIMINARY;	PRT;	403 AA.
AC	043664;			
DT	01-JUN-1998 (TREMBLrel. 06, Created)			
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)			
DE	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	ORPHAN G. PROTEIN-COUPLED RECEPTOR.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxId=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Tan C.T., McKee K.K., Palyha O.C., Feighner S.D., Hreniuk D.L.,			
RA	Smith R.G., Howard A.D.;			
RL	Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.;			
DR	EMBL; AF044601; AAC02680.1;			
DR	EMBL; AF044600; AAC02680.1; JOINED.			
DR	InterPro; IPR000276; GPCR_Rhodopsn.			
DR	Pfam; PF00001; 7tm_1; 1.			
DR	PRINTS; PR00237; GPCRHOOPSP.			
DR	PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.			
DR	PROSITE; PSS0262; G_PROTEIN_RECEP_F1_2; 1.			
SO	SEQUENCE 403 AA; 44980 MW; CFPD5F1A961CCBD9 CRC64;			

Query Match	45.9%	Score 952.5;	DB 4;	Length 403;
Best Local Similarity	47.9%;	Pred. No. 1.9e-05;		
Matches 183;	Conservative 60;	Mismatches 94;	Indels 45;	Gaps 5;

[illegible]

RESULT 8
 ID 09JUT5 PRELIMINARY: PRT: 412 AA.
 AC 09JUT5:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE G PROTEIN-COUPLED RECEPTOR FM-3.
 GN FM-3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-20347213; Pubmed-10783389;
 RA Fujii R., Hosoya M., Fukusumi S., Kawamata Y., Habata Y., Hinuma S.,
 Onda H., Nishimura O., Fujino M.;
 RT Identification of neuropeptide U as the cognate ligand of the orphan G
 protein-coupled receptor FM-3.
 RL J. Biol. Chem. 275:21068-21074(2000).
 DR EMBL; AB038649; BAA9387.1; -.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECPE_FL_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECPE_FL_2; 1.
 KM Receptor.
 SO SEQUENCE 412 AA; 46784 MW; 2CAB56299E2EA7B8 CRC64;

Query Match 44.1%; Score 916.5; DB 11; Length 412;
 Best Local Similarity 49.2%; Pred. No. 1,1e-62;
 Matches 177; Conservative 56; Mismatches 100; Indels 27; Gaps 4;

QY 4 LEMASWIDPL-MKYLNSTEEXLAHLGPKRSDLSLPVSAVALIFLVGMGNLVCNVI 62
 DB LEDLNLTHEDLRLKYL-----GPOVKQFLPICVTYLLIFVVGTLGNGLTCYI 72
 QY 63 VRHQTLKPTNYLFLSLAVSDLLVLLGMPLEIYEMHNYFPLFGVGCYFTALFETVC 122
 DB 73 LROKMHPTNFFYFLSLAVSDLLVLLGMPLEIYEMHNYFPLFGVGCYFTALFETVC 132
 QY 123 FASISVTTSVERVYALVHPRAKLESTRRALRIILSVMSFVSFSLPTISIGIKRPO 182
 DB 133 LASVANTALSERVYAVVHPLAKSVMTKTRVRRMLGIWYFALFSLPTSLHGLSPL 192
 QY 183 HFPNGSVPGSATCTVTRPMVYNIIOATSFYILPMTLISVLYLMLGRRLKRDSELT 241
 DB 193 YVPCGAPVDSTCTLVBPQFFYKLVIOITLLFCLPMTVTSVLYLLGLRLRRRLML 252
 QY 242 -----EANKAVANIHPRSKSVTKMLFVLVFAICMTPPHVDRLFFSEFVEE 288
 DB 293 QEEVKGRIISAARQASHSISIQLRDRERQVTKMLIALVVFCTCWVPEHADRMLSMVSH 312
 QY 289 WTESLAIVFNLIHVYGFVFLYSANVPITVNLSSRFRPAFRNVSPICKCHPRHPO 348
 DB 313 WTDGLRLAFOSVYHLASGVFLYLGSAANPELVNLMSTRERESRETTGLSTRCH-RHOPR 371

RESULT 9
 ID 055040 PRELIMINARY: PRT: 405 AA.
 AC 055040:
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ORPHAN G PROTEIN-COUPLED RECEPTOR.
 GN GPR66.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McKee K.K., Tan C.T., Liu J., Palyha O.C., Feighner S.D.,
 Hreniuk D.L., Smith R.G., Howard A.D.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF044602; AAC02681.1; -.
 DR MGD: MGI:1341898; Gpr66.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECPE_FL_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECPE_FL_2; 1.
 SO SEQUENCE 405 AA; 45609 MW; F1BA493D3DB81F34 CRC64;

Query Match 40.5%; Score 840; DB 11; Length 405;
 Best Local Similarity 46.5%; Pred. No. 8.3e-57;
 Matches 166; Conservative 58; Mismatches 115; Indels 18; Gaps 4;

QY 9 WHDPLMKYLNSTEEYL-AHLGPKRSDLSLPVSAVALIFLVGMGNLVCNVIYRROT 67
 DB 10 WPKQP--EDLNLTHEDLRLKYLGPQOMKQFVPICVTYLLIFVVGTLGNGLTCYIIRNKT 67
 QY 68 LKPTNYLFLSLAVSDLLVLLGMPLEIYEMHNYFPLFGVGCYFTALFETVCASIL 127
 DB 68 MKPTNFFYFLSLAVSDMLVLLGMPLEIYEMQONTFFQLGASACYRILLETVCASVL 127
 QY 128 SVTTSVERVYALVHPRAKLESTRRALRIILSVMSFVSFSLPTISIGIKRPOHPPNG 187
 DB 128 NVTALSERVYAVVHPLAKSVMTKTRVRRMLGIWYFALFSLPTSLHGLSPLPCR 187
 QY 188 SSVPGATCTVTRPMVYNIIOATSFYILPMTLISVLYLMLGRRLKRDSELT----- 241
 DB 188 GPVDSALISGLVPMDFYLVLTALFCLPMTVTSVLYLLGLRLRRRLLOVEVK 247
 QY 242 -----EANKAVANIHPRSKSVTKMLFVLVFAICMTPPHVDRLFFSEFV-EEFTES 292
 DB 248 GKRTAQTGTSRHRIDLODRGRQVTKMLFALVYVVGICMAPRHADRIMSLVYGHSTG 307
 QY 293 LAAVFNLIHVSGVFYLSANVPITVNLSSRFRPAFRNVSPICKCHPRHPOG 349
 DB 308 LHLAYOCVHIASGIFFYLGSAANPVLVYLSMSTRFETFOALGLGTCQCHRRQPYHG 364

RESULT 10
 ID 09VFN4 PRELIMINARY: PRT: 428 AA.
 AC 09VFN4:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CG9918 PROTEIN.
 GN CG9918.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriodes; Drosophilidae; Drosophila.
 NC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE-20196006; Pubmed-10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.-H.C., Blazer J.R.G., Chapple M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agdayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibeigbam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., Moberly D.,
 RA Merkulyov G., Mlshina N.V., Moberly C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K.A., Nusskern D.R., Pacle J.M.,
 RA Palczolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000)
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC EMBL: AE003703; AAF55016.1; -
 DR FlyBase: FBgn00384201; CG918.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm.1; 1
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PROSITE: PS00237; G-PROTEIN_RECEP_FL_1; 1.
 DR PROSITE: PS00262; G-PROTEIN_RECEP_FL_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Transmembrane.
 SO SEQUENCE 428 AA; 47198 MW; 9DDA41B93046BD7 CRC64;

Query Match 27.2%; Score 564.5; DB 5; Length 428;
 Best Local Similarity 32.2%; Pred. No. 1.2e-35;
 Matches 118; Conservative 78; Mismatches 104; Indels 67; Gaps 6;

25 LAHLGPKRSDLS--LPVSAYALFLVGVGMLVCMVVRHOTLTKPTNYLFLSAVS 82
 6 MSHDLGPPRPDPLAIVIPVYVYSLRITGVGNISCTIVAKKNSMTATNYLFLSLAIS 65
 83 DLVLVLLGMPLEIYEMWVNPFLGPGVCYFKTALFETVCFASILSVTVSVRYVAIVH 142
 66 DFLLLLSVGPQDEYSYIMSKRPYVFEVETICRGLLAETSANAVLTATVEYNTACH 125
 143 PPRAKLESTRRRRLRLISLWSFSVPSLPNTSHIGIKFOHPFGNSVPGSATCTYTKPM 202
 126 PFLGQMSKLSRLRIIVLWMAIVTAIPQAAQFGI--EHY-----SCVEDCGIVRVI 177
 203 WYVNLIIQATSPFLYILPMTLSLVLYMGLRLKRDSELSA----- 243
 178 VKHS--FOLSTFLFLAPMSIILVLLYLGVLHLYGVGPASVARROOLKSVPSDTIL 235
 244 -----NKAVINHRPSRKSYSYTKMLFVLVLEFAI 271
 236 YRGGSGTAMSFNGGSGAGTACLMGSSGAGSLSVRRGLRHLHYGRRLRLRLVAVVCFEL 295
 272 CWTPEVVDRLFSFV---EEMTESLAAEVNLHVVSGCFEYSSAVNPITLYNLSRRR 327
 296 CMAPFAORLAIYAPARGAKLRDQHEFYVTVMYVSGVLYSTCINPLLYNIMSHKFR 355
 328 AAFRNVV 334
 356 EAFKAVL 362

RESULT 11
 ID 09VFW6 PRELIMINARY; PRT; 595 AA.
 AC 09VFW6;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CG8795 PROTEIN.
 GN CG8795
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; *Drosophila*.
 OC NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amandides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrial J.F., Agdanyani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bottler P.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bottler P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibeigbam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulyov G., Mlshina N.V., Moberly C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palczolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC EMBL: AE003699; AAF54929.1; -
 DR FlyBase: FBgn0038139; CG8795.
 DR InterPro: IPR000923; Copper_blue1.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm.1; 1
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PROSITE: PS00196; COPPER_BLUE; UNKNOWN.1.
 DR PROSITE: PS00237; G-PROTEIN_RECEP_FL_1; 1.
 DR PROSITE: PS00262; G-PROTEIN_RECEP_FL_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Transmembrane.
 SO SEQUENCE 595 AA; 65091 MW; 3EBE183BADF5471E CRC64;

Query Match 26.3%; Score 546.5; DB 5; Length 595;

Best Local Similarity 35.1%; Pred. No. 4e-34;
Matches 120; Conservative 68; Mismatches 115; Indels 39; Gaps 9;

QY 14 LKLYNSTEYLAHLCGPRSDSLP-----VSVAIALFLVGVGNLLVCM 60
DB 26 LQVLLISADNLSLQGLEPELPTVPMPLSLATLSVGLFVAGNLTCTI 85
QY 61 VIVRHOTLPTNYLFSLAVSDLLVLLGMPLEIYEMMH--VYPLFGVGVGKRALF 118
DB 86 VTSRNNFMHTATNFYLFNLAISMDILCSMPDGLNLMHPDNP--FSDSITLESVLS 143
QY 119 ETVCFASISLVTTVSERYVAIYPRAKLESTRRALRLSLWSSEVSFSLPNTSING 178
DB 144 EFMANATVLTIFAFTERYIAICHPROHTMSKLSRAVKRTFAIMWLAALLALP---QA 199
QY 179 IKFQHPNNSVPGSATCIVTKPMWYNLIQTSFLYILPMTLSLVLYLGLRLKRD 238
DB 200 IGFSVVMQGM---GTSCMKNDFAH--VFAVSGFLFEGGPMATCVLVVLGVKLRKS 253
QY 239 ESLEA---NKVAVNIHRPSKSVTKMLFVLVLFVLCMPFHVDRLEFSF-----VEEM 289
DB 254 RLQALPRCIVDNRGISNOTRIYRLVAVAVAFICMAFPHQRLMAVAGTSGIESQM 313
QY 290 TESLAAVFLNIHVYGVFFYLLSSAVNPITYNLLSRFRRAFR 331
DB 314 FND---VFSIDYTSGLVFLSTCINPLLYNIMSHKFRFAFK 352

RESULT 12
QYVFWS PRELIMINARY; PRT; 660 AA.

AC 09VFWS; 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE CG8784 PROTEIN.
GN CG8784.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidae; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN
RP
RC
RA MEDLINE-20196006; PubMed=10731132;
RA STRAIN-BERKELEY;
RA Adams M.D., Celinkler S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA Abri J.F., Abdayani A., An H.-J., Andrews-Planckoch C., Baldwin D.,
RA Bailey R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Danke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeigbam C.,
RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasno P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Modarity C., Morris J., Mostreft A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skrupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wattaman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC 1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC 1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: AE003699; AAF5430.1; -
DR FlyBase: FBgn0038140; CG8784.
DR InterPro: IPR000923; Copper_D1uel.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSN.
DR PROSITE: PS00196; COPPER_BLUE; UNKNOWN_1.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ SPROUCE 660 AA; 72277 Mw; C52D1EAT4ACD439 CRC64;

Query Match 25.3%; Score 525.5; DB 5; Length 660;
Best Local Similarity 29.7%; Pred. No. 1.8e-32;
Matches 127; Conservative 74; Mismatches 153; Indels 73; Gaps 9;

QY 18 LNSTEYLAHLCG-----PKRSDSLP--PVSAVALFLVGVGNLLVCMV 61
DB 72 LNTTELSNLGSGTNGTNTMAADSPVDESLTRALTYCAVLAIVAGVGLTLCIV 131
QY 62 IYRHQTKPTNYLYFLSLAVSDLLVLLGMPLEIYEMMHNPFLGPGVGYFKTALFETV 121
DB 132 ISRNFMHTATNFYLFNLAISMDILVSGIPOELNLMYDPDMPFTDAMCIMGSVLSMA 191
QY 122 CPASISLVTTVSERYVAIYHPRAKLESTRRALRLSLWSSEVSFSLPNTSING 181
DB 192 ANAVLITITFAFTERYIAICHPROHTMSKLSRAVKRTFAIMWLAALLALPQAMQFSVY 251
QY 182 QHPFGSSVPGSATCIVTKPMWYNLIQTSFLYILPMTLSLVLYLGLRLKRD--- 238
DB 252 QN---EGTS-----CTENDEYAH--VFANSGFLFEGGPMATCVLVVLGVKLRRL 301
QY 239 ESLEANKVAVNIHRPSKSVTKMLFVLVLFVLCMPFHVDRLEFSF-----VEEMTES 292
DB 302 QSLPRRTFDANRGLNAGGRIYRLVAVAVAFICMAFPHQRLMAVAGNLTINIGISRDA 361
QY 293 LAAVENLIHVYGVFFYLLSSAVNPITYNLLSRRAAFRNVSPTCKWCCHPRRPG--- 349
DB 362 FNDFFRLDYTSGLVFLSTCINPLLYNIMSHKFRFAFKITLTFQGLARNHHHQSQNH 421
QY 350 -----PPAKIILFTCHLVELPEDAGORPGSSI 380
DB 422 QHNTSALLRONGSKRLQPCASVNNNALPEYGVTRVQFRC-----RDANHQLSDSI 475
QY 381 HNTNLT 387
DB 476 RTTITTT 482

RESULT 13
ID 017239 PRELIMINARY; PRT; 418 AA.
AC 017239; 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE K10B4.4 PROTEIN.
GN K10B4.4.
OS Caenorhabditis elegans.

SEQUENCE FROM N.A.
 RP Palyha O.C., Tan C., McKee K.K., Felghner S.D., Hrenluk D.L.,
 RA Smith R.G., Howard A.D.;
 RT "Orphan G protein-coupled receptor from teleost fish Spheroides
 nelpelus related to growth hormone secretagogue receptor.";
 RL Submitted (AUG-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF082209; AAC3472.1; .
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR000524; HTH_GntR.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECCEP_F1_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECCEP_F1_2; 1.
 DR PROSITE: PS00043; HTH_GNTR_FAMILY; UNKNOWN_1.
 SO SEQUENCE 374 AA; 42324 MW; 2CE9304F004C7A16 CRC64;

Query Match 22.6%; Score 470; DB 13; Length 374;
 Best Local Similarity 33.1%; Pred. No. 1,8e-28;

Matches 118; Conservative 73; Mismatches 117; Indels 48; Gaps 10;

```

OY 2 GLEENASNIHDPKMYLNSTEEYLAHLGCPKRSLSLP---VSAVALIFLVGMGILL 57
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 6 GLSPNCSEW-----EGSHNGTAGLELPPLNYYSIPLAVITVACTVLFVGVGNVM 56

OY 58 VCMVIVRHOTLKTPTNYVLFSLAVSDLLVLLGMPLEIYEMMHNPFLFGPVGCYEKTAL 117
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 57 TLIVSRTRDMHTTNILCSMAVSDFLFTVC-MPLDLYRMNRYRPMRFGDALCKLQFV 115

OY 118 FETVCFASILSVTVSVERYVAIVHPFRAKLESTRRALRLILSVSFSVSFLPNTSIH 177
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 116 SESCTYSTLCITALSEYRYLAICFPLRAKALVTKRRVALLILLTVSLSGPVFVMV 175

OY 178 GIKFQH--FPNGSSVPGSA-----TCTVTK-----PMVYNLIITQATSLFL 216
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 176 GVEKDSIMEPNSSDLNESSMPLAVDTRECRMTQYAVESGLMEAMVM-----SSVF 227

OY 217 YLLPMTLISVLYLGLR--LKRDESLKANKVAVNIRPSRKSVTKMLFVLVLFALICWT 274
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 228 FPMVFCLTVLGLGRRLMLRHRETTINSRYA--YRDKSNKOTIKMLVYVVLAFVLCML 285

OY 275 PFHVDR-LFFSFVEEMTESLAAVFNLIHVSGVFYLSAVNPDIYNLRSRRRAA 329
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 286 PFHVGRVYLQFRSLDAPSLSLSEYCSLSVYVLFYLSAINDPILYNTMSWKYRGA 341

```

Search completed: April 22, 2002, 14:19:19

Job time: 337 sec

THIS PAGE BLANK (USPTO)

[illegible]

```

FT Domain 202..227
FT /Label= Transmembrane_domain
FT 259..278
FT /Label= Transmembrane_domain
FT 292..322
FT /Label= Transmembrane_domain
PN W0200144297-A1.
XX
XX
PD 21-JUN-2001.
XX
XX
PF 13-DEC-2000; 2000WO-US33787.
XX
XX
PR 17-DEC-1999; 99US-0466435.
PR 25-APR-2000; 2000US-0558099.
PR 30-JUN-2000; 2000US-0609146.
XX
XX
PA (SYNA-) SYNAPTIC PHARM CORP.
PI Bonini JA, Lerman GS, Quan Y, Ogozalek K;
XX
XX
DR WPI: 2001-390240/41.
DR N-PSDB; AAD08024.
XX
XX
PT A purified mammalian SNORF62 or SNORF72 receptor protein for
PT identification of compounds to treat e.g. inflammation, arthritis,
PT autoimmune diseases, transplant rejection, AIDS, cancer -
XX
XX
PS Claim 15; Fig 15; 256pp; English.
XX
XX
CC The invention relates to human G-protein coupled receptors, SNORF62 and
CC SNORF72 and their corresponding cDNA molecules. SNORF62 and SNORF72
CC receptors are specific for neuropeptide U (NMU) neuropeptides, hence they
CC are also known as NMU receptors. The agonist and antagonist of NMU
CC receptors are useful for treating an abnormality in a subject that is
CC alleviated by decreasing or increasing the activity of NMU receptor.
CC The NMU receptor serves as a valuable tool for designing drugs which are
CC useful for treating various pathological conditions such as
CC inflammation, arthritis, autoimmune diseases, transplant rejection,
CC graft vs host disease, bacterial, fungal, protozoan and viral infections,
CC septicemia, AIDS, pain, psychotic and neurological disorders, including
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, neuromotor disorders, respiratory disorders, asthma,
CC eating/body weight disorders including obesity, bulimia, diabetes,
CC anorexia, nausea, hypertension, hypotension, vascular and cardiovascular
CC disorders, ischemia, stroke, cancers, sexual disorders, circadian
CC disorders, renal disorders, bone diseases including osteoporosis, benign
CC prostatic hypertrophy, gastrointestinal disorders, nasal congestion,
CC dermatological disorders such as psoriasis, allergies, Parkinson's
CC disease, Alzheimer's disease, acute heart failure, angina disorders,
CC delirium and dyskinetic diseases such as Huntington's disease. They can also be
CC used to regulate steroid hormone disorders, epinephrine release
CC disorders, electrolyte balance disorders, endocrine disorders, memory
CC disorders, somatosensory disorders, metabolic disorders, behavioural
CC disorders, drug addiction, migraine, Addison's disease, Cushing's
CC disease, prevent miscarriage, induce labour or to treat dysmenorrhea.
CC The present sequence is rat G-protein coupled receptor, SNORF72 related
CC to the invention.
XX
XX
SQ Sequence 395 AA:

```

```

Query Match 100.0%; Score 2076; DB 22; Length 395;
Best Local Similarity 100.0%; Pred. No. 4, 9e-227;
Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MGRLENSWTHDPLMKYLNSTEEYLAHLGPKRSDDLSPVSVAALFVLGVGNLLVCM 60
DB 1 mglenswshdpkmkylnsteeylahlgpkrsdslspvsveyallflvgvgnllvcm 60
QY 61 VIVRHQTLPYNYFLSLAVSDLLVLLGMPLEIYEMHNPFLFGPVGCYKTALEET 120
DB 61 vlvrhqtlpnyflslavsdllvllgmpleyemhnpflfgpvgykftalet 120

```

```

QY 121 VCFASILSVTVSVRYVAIVHPRAKLESTRRRALRLSLWSESVFSLPNTSINGIK 180
DB 121 vcfasilsvtvsveryaivhpraklestrrralrlslwsefvfslpntsingik 180
QY 181 FOHPNGSSVPGSATCTVTKRMVWYNLIQATSEFLYFLPMTLISVLVYLMGLRKDES 240
DB 181 fohpngssvpgsatctvtkrmvwnyliqatseflfyflpmtlislvlvylmglrkdes 240
QY 241 LEANKVAVNTHRPKRKSVTKMLFVLVFAICWTFPHVDRLFFSFVEEMTESLAAVNLI 300
DB 241 leankvavnthrpkrksvtkmlfvlvfaicwtfphvdrlffsfveemeslaavnli 300
QY 301 HVSQGFVFLSSAVNPITVNLSSRRFAAFRNVSPTCKMCHPRRPOGPAQKIIFLTE 360
DB 301 hvsgfvflssavnpitvnlssrrfaafrnvsptckmchprpogpaqkiiflte 360
QY 361 CHLVELTEDAGPPGQSSIHNTNLTPAPCAGEVP 395
DB 361 chlveltedagpgfgqssihntnltpapcagevp 395

```

```

RESULT 2
AAG63367
ID AAG63367 standard; Protein: 395 AA.
XX
AC AAG63367;
XX
DT 15-OCT-2001 (first entry)
XX
XX Amino acid sequence of a rat TGR-1 protein.
XX
XX TGR-1; neuropeptide U; hypertension; stress disease.
XX
XX Rattus sp.
XX
XX W0200157524-A1.
XX
XX 09-AUG-2001.
XX
XX 02-FEB-2001; 2001WO-JP00746.
XX
XX 04-FEB-2000; 2000JP-0032773.
XX
XX 24-FEB-2000; 2000JP-0052252.
XX
XX 30-MAR-2000; 2000JP-0097896.
XX
XX 19-JUN-2000; 2000JP-0187536.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;
PI Okubo S;
XX
DR WPI: 2001-488917/53.
DR N-PSDB; AAA43078.
XX
PT Identifying predicted or actual structures of two or more members of a
PT chemical or physical library by mass spectrometry comprising
PT correlating molecular mass measurements of two or more members with a
PT shared chemical history -
XX
XX
PS Claim 1; Page 89-91; 95pp; Japanese.
XX
XX The present sequence represents a rat TGR-1 protein. The specification
XX describes a method of screening a compound, which is capable of binding
XX properties of neuropeptide U to TGR-1. The method is useful for screening
XX preventatives and remedies for hypertension, stress diseases, etc..
XX TGR-1 antagonists are also useful for treating the same diseases.
XX
XX
SQ Sequence 395 AA:

```

```

Query Match 99.4%; Score 2063; DB 22; Length 395;
Best Local Similarity 99.2%; Pred. No. 1, 5e-225;

```

Matches 392; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
OY 1 MGKLENAWIMHDPKMYNSTEETVLAHICGPKRSDLSLPVSAVALFLFYGMGNLVCNM 60
DB 1 mglknaawimdpkmylnsteevylahicgpkrsdlsipvsavallflfygmgnlvcnm 60
OY 61 VVRRHQTPTPNYYLFSLAVSDLVLLGMPLEIYEMHNPFLFGPYGCFKTALEET 120
DB 61 vvrhqtptpnnylflslavsdllvllgmpleyemhnpflfpgpycfktalet 120
OY 61 vlvrtqltkrpnnylflslavsdllvllgmpleyemhnpflfpgpycfktalet 120
DB 61 vlvrtqltkrpnnylflslavsdllvllgmpleyemhnpflfpgpycfktalet 120
OY 121 VCFASILTSTVTVSVRYVAIVHPRAKLESTRRALRLISLWFSQVFSLPNTSIHGK 180
DB 121 vcfasiltstvtvsryvaivhpraklestrrralrlislwfsqvfsfslpntshgk 180
OY 181 FOHFNGSSVPSATCTYTKRPMVYNLLIQTATSPLEIYILPMLISLVLYMLRLKRDS 240
DB 181 fohfngssvpsatctytkrpmvynlliqtatspleyilpmlislvlymlrlkrds 240
OY 241 LEANKVAVNIHPRSRKSVTKMLFVLVFAICWTPFHVDRLFFSEVENTESLAAVFNLI 300
DB 241 leankvavnihparskvtkmlfvlvfaicwtpfhvdrlffseventeslaavfnli 300
OY 301 HVSQVFFPLSSAVNPITIIYNLISRRFRAFRNVSPCTCKMCHPRHROGPPAQKIIFLTE 360
DB 301 hvsgvffplyssavnpiliiynlisrrfrafrrnvspctckmchprhroqppaqkiiflte 360
OY 361 CHLVLETDAGPQPGOSSIHNTNLTTAPCAGEVP 395
DB 361 chlveltedagpqfpgqsalhntnlmapcagevp 395
```

RESULT 3

AAB67803 ID AAB67803 standard: Protein: 412 AA.

AC AAB67803:

DT 29-JUN-2001 (first entry)

DE Amino acid sequence of G-protein coupled receptor IGSA short version.

Human; G-protein coupled receptor; IGSA; IGSA4; IGSA4B; schizophrenia;
 nervous system disorder; psychiatric disorder; Parkinson's disease;
 episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
 cardiovascular disease; heart failure; angina pectoris; obesity; anorexia;
 motility disorder; myocardial infarction; hypertension; dyslipidemia;
 gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
 inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
 gynecological disorder.

OS Homo sapiens.

PN WO200125269-A2.

PD 12-APR-2001.

PE 25-SEP-2000; 2000WO-EP09584.

PR 24-SEP-1999; 99EP-0203140.

PR 24-SEP-1999; 99NL-1013140.

PR 28-JUL-2000; 2000EP-0202683.

PR 31-JUL-2000; 2000US-0222047.

PA (SOLV) SOLVAY PHARM BV.

PI Deleersnyder W, Berger C, Loeken C, Nys G, Venema J;

DR MPI; 2001-273568/28.

DR N-PSDB; AAF80323.

PT New G-protein coupled receptors and the polynucleotides encoding them,
 useful for preventing, ameliorating or correcting nervous system
 disorders, cardiovascular diseases, dyslipidemias, inflammations, pain

PT or cancers -
 XX Claim 18; Page 81-82; 102pp; English.

The present sequence represents the short version of a human G-protein coupled receptor designated IGSA4. IGSA4 exists in two polymorphic forms, CC IGSA4 and IGSA4B. The IGSA4 receptors and IGSA4 polynucleotides are useful for preventing, ameliorating or correcting dysfunctions or diseases. These diseases include peripheral nervous system, psychiatric and central nervous system disorders (e.g. schizophrenia, episodic paroxysmal anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or stroke), cardiovascular diseases (e.g. heart failure, angina pectoris, myocardial infarction or hypertension), dyslipidemias, obesity, anorexia, gastrointestinal disorders (e.g. inflammatory bowel disease or motility disorders), osteoporosis, inflammations, infections (e.g. bacterial, fungal, protozoan or viral), pain, cancers, immune disorders, allergies, sepsis or gynecological disorders. Agonists or antagonists of IGSA4 are effective with regard to disorders of the nervous system, including the central and peripheral nervous systems, disorders of the gastrointestinal system, cardiovascular system, skeletal muscle, thyroid, lung or genitourinary system, or immunological disease. The IGSA4 polynucleotides are useful as diagnostic reagents for detecting under-expression, overexpression or altered expression of IGSA4.

Sequence 412 AA;

Query Match 79.5%; Score 1650.5; DB 22; Length 412;
 Best Local Similarity 79.4%; Pred. No. 1,le-178;
 Matches 312; Conservative 36; Mismatches 40; Indels 5; Gaps 1;

```
OY 1 MGKLENAWIMH-----DPLMKYLNSTEEVLAHICGPKRSDLSLPVSAVALFLFYGMGN 55
DB 1 mglknaawimhpyqkledpqklnsteevylahicgpkrsdlsipvsavallflfygmgn 60
OY 56 LVCWVYVNRHQTPTPNYYLFSLAVSDLVLLGMPLEIYEMHNPFLFGPYGCFKTALEET 115
DB 61 vlvclvlylqgkamkprnylflslavsdllvllgmpleyemhnpflfpgpycfktalet 120
OY 116 ALFETVCPASITSVTVSVRYVAIVHPRAKLESTRRALRLISLWFSQVFSLPNTSIHGK 175
DB 121 alfetvcasitstvtvsryvaivhpraklestrrralrlislwfsqvfsfslpntshgk 180
OY 176 HGKIFRHPNGSSVPSATCTYTKRPMVYNLLIQTATSPLEIYILPMLISLVLYMLRLKRDS 235
DB 181 hgkifrhpnssvpsatctytkrpmvynlliqtatspleyilpmlislvlymlrlkrds 240
OY 236 KRDSELEANKVAVNIHPRSRKSVTKMLFVLVFAICWTPFHVDRLFFSEVENTESLA 295
DB 241 krdseleankvavnihparskvtkmlfvlvfaicwtpfhvdrlffseventesla 300
OY 296 VFNLIHVSQVFFPLSSAVNPITIIYNLISRRFRAFRNVSPCTCKMCHPRHROGPPAQKI 355
DB 301 vfnlihvsgvffplyssavnpiliiynlisrrfrafrrnvspctckmchprhroqppaqki 360
OY 356 IFLTECHLVLETDAGPQPGOSSIHNTNLTTA 388
DB 361 ifltechlvletdagpqfpgqsalhntnlmapcagevp 393
```

RESULT 4

AAV71296 ID AAV71296 standard: Protein: 415 AA.

AC AAV71296:

DT 02-NOV-2000 (first entry)

DE Human orphan G protein-coupled receptor hRUP6.

Human; orphan G protein-coupled receptor; GPCR; hRUP6; drug screening;
 transmembrane receptor; signal cascade.

```
OS Homo sapiens.
XX WO200031258-A2.
XX
XX 02-JUN-2000.
XX
XX 13-OCT-1999. 99WO-US23687.
XX
XX 20-NOV-1998. 98US-0109213.
XX 16-FEB-1999. 99US-0120416.
XX 26-FEB-1999. 99US-0121852.
XX 12-MAR-1999. 99US-0123946.
XX 12-MAR-1999. 99US-0123949.
XX 28-MAY-1999. 99US-0136436.
XX 28-MAY-1999. 99US-0136437.
XX 28-MAY-1999. 99US-0136439.
XX 28-MAY-1999. 99US-0136567.
XX 28-MAY-1999. 99US-0137127.
XX 28-MAY-1999. 99US-0137131.
XX 29-JUN-1999. 99US-0141448.
XX 29-SEP-1999. 99US-0156555.
XX 29-SEP-1999. 99US-0156633.
XX 29-SEP-1999. 99US-0156634.
XX 29-SEP-1999. 99US-0156653.
XX 01-OCT-1999. 99US-0157280.
XX 01-OCT-1999. 99US-0157281.
XX 01-OCT-1999. 99US-0157282.
XX 01-OCT-1999. 99US-0157293.
XX 01-OCT-1999. 99US-0157294.
XX 12-OCT-1999. 99US-0416760.
XX 12-OCT-1999. 99US-0417044.
XX
XX (AREN-) ARENA PHARM INC.
XX
XX PI Chen R., Dang HT, Llaw CW, Lin I.
XX
XX WPI: 2000-40068/34.
XX DR N-PSDB; AAD01123.
XX
XX PT Novel human orphan G protein-coupled receptors and the encoding cDNAs
XX for use in the identification of G protein-coupled receptor agonists -
XX
XX Claim 22: Page 57-59; 102pp; English.
XX
XX PS The present amino acid sequence is the hrup6, an endogenous human
XX CC orphan G protein-coupled receptor (GPCR). The full length hrup6 cDNA
XX CC was cloned by RT-PCR using human thymus cDNA as template.
XX CC The orphan GPCR of the invention, like all GPCRs has seven transmembrane
XX CC alpha helices with an extracellular N-terminus and an intracellular
XX CC C-terminus. However, no endogenous ligands has yet been identified for
XX CC the proteins of the invention. The orphan GPCRs may be used in the
XX CC identification of their endogenous ligands, and to screen potential GPCR
XX CC agonists and antagonists for use as pharmaceutical agents. The proteins
XX CC may also be used in the study of GPCR-mediated signalling cascades, and
XX CC to elucidate their precise role in normal and diseased human conditions.
XX CC Nucleic acid encoding human orphan GPCRs may be used for tissue
XX CC localisation expression analysis to provide information about their
XX CC function in healthy and pathological states.
XX
XX Sequence 415 AA:

Query Match 79.5%; Score 1650.5; DB 21; Length 415;
Best Local Similarity 79.4%; Pred. No. 1,1e-178;
Matches 312; Conservative 36; Mismatches 40; Indels 5; Gaps 1;

QY 1 MGLKLENAWTH-----DPLMKYLNSTEEYLAHLGPKRSDLSLPVAVAYALFLVGMGN 55
DB 4 meklgnasvlygqkledpfqkhnsteeylaficgprtrshffilpvasvvyplfvvgvign 63
QY 56 LLYCMTIVIRHOTLKPTPNYLYSLAVSDLLVLLGMPLELYEMWNNHYRFLFGVGCYFKT 115
DB 64 vlycvllylqhqamkrcptnylylsjavsdlivlllmplyemwvnyrpfifgvgcyfkt 123
```

```
QY 116 ALPEYCFASILSVTTVSVERVVAIVHPRAKLESTRRLRILSLWSPSVFSLPNTS 175
DB 124 alfevcfasilsvttvsvervvaillhpfraklqstrrralrlglvwgsvfslpntcs 183
QY 176 HGIRKQHPNPNGSSVPGSATCTVTKRPMWYNLIQATSFLEYLLPMTLISVLYLGLRL 235
DB 184 hbgirkfhypngslvpgsatctvtkrpmwlynliqatstfleyllpmtvlsvlylmalr 243
QY 236 KRDESLKANKVAVNIHRPSKSVTKMLFVLYLVFAICWPFHVDRLFFSVFEWTESLAA 295
DB 244 krdksleadeqnanlqrcpksvtnkmlfvlylvfaicwaphldrlffsfvewseslaa 303
QY 296 VFNLHVSGVFYLLSAVNPITYNLSRFRRAFRNVVSPCTCKWCHPRHRPGPPAOKI 355
DB 304 vfnlvhvsqvfyllyssavnpilynllstrfgaefgnvlsfkkqwhsqhdpqjppaqn 363
QY 356 IFLECHLVELTEDAGPQPPGQSSIHNTNLTTA 388
DB 364 iflechveltedlqpfqgssmshlptra 396

RESULT 5
AAB02830
ID AAB02830 standard; Protein: 415 AA.
XX
XX AAB02830;
AC
XX 22-AUG-2000 (first entry)
DT
XX
XX Human G protein coupled receptor hrup6 protein SEQ ID NO:12.
DE
XX
XX Human: G protein coupled receptor; GPCR; transmembrane receptor;
XX identification; agonist; screening; therapeutic; pharmaceutical;
XX mutant.
XX
XX OS Homo sapiens.
XX
XX PN WO200022131-A2.
XX
XX PD 20-APR-2000.
XX
XX PF 13-OCT-1999; 99WO-US24065.
XX
XX PR 13-OCT-1998; 98US-010496.
XX PR 12-NOV-1998; 98US-0108029.
XX PR 20-NOV-1998; 98US-0109213.
XX PR 27-NOV-1998; 98US-0110060.
XX PR 16-FEB-1999; 99US-0120416.
XX PR 26-FEB-1999; 99US-0121852.
XX PR 12-MAR-1999; 99US-0123944.
XX PR 12-MAR-1999; 99US-0123945.
XX PR 12-MAR-1999; 99US-0123946.
XX PR 12-MAR-1999; 99US-0123948.
XX PR 12-MAR-1999; 99US-0123949.
XX PR 12-MAR-1999; 99US-0123951.
XX PR 28-MAY-1999; 99US-0136436.
XX PR 28-MAY-1999; 99US-0136437.
XX PR 28-MAY-1999; 99US-0136439.
XX PR 28-MAY-1999; 99US-0137127.
XX PR 28-MAY-1999; 99US-0137131.
XX PR 28-MAY-1999; 99US-0137567.
XX PR 30-JUN-1999; 99US-0141448.
XX PR 27-AUG-1999; 99US-0151114.
XX PR 03-SEP-1999; 99US-0152524.
XX PR 29-SEP-1999; 99US-0156633.
XX PR 29-SEP-1999; 99US-0156655.
XX PR 29-SEP-1999; 99US-0156634.
XX
XX (AREN-) ARENA PHARM INC.
XX
XX PI Behn DP, Lehmann-Brulsma K, Chalmers DT, Chen R, Dang HT;
PI Gore M, Llaw CW, Lin I, Lowitz K, White C;
```

XX WPI: 2000-317986/27.
 DR N-PSDB: AAA46022.
 XX
 PT Non-endogenous, human G protein-coupled receptors for screening
 XX receptor, inverse or partial agonists useful as therapeutic agents
 XX
 PS Example 1: Page 86-88; 187pp; English.
 XX
 CC The present invention describes transmembrane receptors, preferably
 CC human G protein coupled receptors (GPCR), for which the endogenous
 CC ligand is unknown (orphan GPCR receptors). More specifically the present
 CC invention relates to non-endogenous, constitutively activated versions
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for
 CC the direct identification of candidate compounds as receptors agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents.
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 CC the exemplification of the present invention.
 CC
 SO Sequence 415 AA:

Query Match 79.5%; Score 1650.5; DB 21; Length 415;
 Best Local Similarity 79.4%; Pred. No. 1.1e-178;
 Matches 312; Conservative 36; Mismatches 40; Indels 5; Gaps 1;

QY 1 MGKLENAWTH-----DPLMKYLNSTEEYLAHLCGPKRSDLSLPVAVAAALFLVGMGN 55
 DB 4 meklqnaswlyqgkledpfqkhnsteeylaficgprtrhfflpvsvvypflfvvgvign 63
 QY 56 LTVCAVIVRHQTLKTPNTNYLFESLAVSDLVLLGMPLEIYEMHNYPLFGVGCYFKT 115
 DB 64 vlvclvllqhgankrptnyllfslavsdllvllgmpleymwrryplflfpvgcyfkt 123
 QY 116 ALFETVCFASISLVTVSVSRVRYAIVHPRAKLESTRRRLRLISLVWSEVSFLPNIS 175
 DB 124 alftvcfasllsltlcvsvryaallhprfkqstrrralrlilglvwgsvflsfpnts 183
 QY 176 IHGIRKQHPNGSSVPGSATCTVTKPMWYNLIQATSEFLYILPMTLSVLVYLGRL 235
 DB 184 ihgikthyfngslvpgsatctvfkpmwlynfllyqtsflfyllpmtvlsvlylmaelrl 243
 QY 236 KRDESLKANKVAVNIHRPSKSVTKMLFVLVLAICWTFPHVDRLFESFVEEMTESLAA 295
 DB 244 kkdksleadeqnanlqprcksvtkmlfvlvlyfalcwtfphldrlffsfveeslaa 303
 QY 296 VFNLHVSGVFYSSAVNPITLYNLSRRFRARFNVVSPCKMCHRRHRRPGAPRAOKI 355
 DB 304 vfnlvhvsqvfylssavnpilnyllstrfgaafrnvsvfkhqwsqndpqlppaqrn 363
 QY 356 IFLTECHLVLELTEDAGPQFGGSSIHNTNLT7A 388
 DB 364 ifltechfveltedlqpfqfgssmshnpa 396

RESULT 6
 AAC63353
 ID AAC63353 standard; Protein; 415 AA.
 XX
 AC AAG63353;
 XX
 DT 15-OCT-2001 (first entry)
 XX
 DE Amino acid sequence of a human TGR-1 protein.
 XX
 KW TGR-1; neuromedin U; hypertension; stress disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157524-A1.
 XX
 PD 09-AUG-2001.
 XX

PF 02-FEB-2001; 2001WO-JP00746.
 XX
 PR 04-FEB-2000; 2000JP-0032773.
 PR 24-FEB-2000; 2000JP-0053252.
 PR 30-MAR-2000; 2000JP-0097896.
 PR 19-JUN-2000; 2000JP-0187536.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;
 PI Okubo S;
 XX
 DR WPI: 2001-488917/53.
 DR N-PSDB: AAA43072.
 XX
 PT Identifying predicted or actual structures of two or more members of a
 PT chemical or physical library by mass spectrometry comprising
 PT correlating molecular mass measurements of two or more members with a
 PT shared chemical history -
 XX
 PS Claim 1: Page 77-79; 95pp; Japanese.
 XX
 CC The present sequence represents a human TGR-1 protein. The specification
 CC describes a method of screening a compound, which is capable of binding
 CC properties of neuromedin U to TGR-1. The method is useful for screening
 CC preventatives and remedies for hypertension, stress diseases, etc..
 CC TGR-1 antagonists are also useful for treating the same diseases.
 CC
 SO Sequence 415 AA:

Query Match 79.5%; Score 1650.5; DB 22; Length 415;
 Best Local Similarity 79.4%; Pred. No. 1.1e-178;
 Matches 312; Conservative 36; Mismatches 40; Indels 5; Gaps 1;

QY 1 MGKLENAWTH-----DPLMKYLNSTEEYLAHLCGPKRSDLSLPVAVAAALFLVGMGN 55
 DB 4 meklqnaswlyqgkledpfqkhnsteeylaficgprtrhfflpvsvvypflfvvgvign 63
 QY 56 LTVCAVIVRHQTLKTPNTNYLFESLAVSDLVLLGMPLEIYEMHNYPLFGVGCYFKT 115
 DB 64 vlvclvllqhgankrptnyllfslavsdllvllgmpleymwrryplflfpvgcyfkt 123
 QY 116 ALFETVCFASISLVTVSVSRVRYAIVHPRAKLESTRRRLRLISLVWSEVSFLPNIS 175
 DB 124 alftvcfasllsltlcvsvryaallhprfkqstrrralrlilglvwgsvflsfpnts 183
 QY 176 IHGIRKQHPNGSSVPGSATCTVTKPMWYNLIQATSEFLYILPMTLSVLVYLGRL 235
 DB 184 ihgikthyfngslvpgsatctvfkpmwlynfllyqtsflfyllpmtvlsvlylmaelrl 243
 QY 236 KRDESLKANKVAVNIHRPSKSVTKMLFVLVLAICWTFPHVDRLFESFVEEMTESLAA 295
 DB 244 kkdksleadeqnanlqprcksvtkmlfvlvlyfalcwtfphldrlffsfveeslaa 303
 QY 296 VFNLHVSGVFYSSAVNPITLYNLSRRFRARFNVVSPCKMCHRRHRRPGAPRAOKI 355
 DB 304 vfnlvhvsqvfylssavnpilnyllstrfgaafrnvsvfkhqwsqndpqlppaqrn 363
 QY 356 IFLTECHLVLELTEDAGPQFGGSSIHNTNLT7A 388
 DB 364 ifltechfveltedlqpfqfgssmshnpa 396

RESULT 7
 AAC64297
 ID AAC64297 standard; Protein; 415 AA.
 XX
 AC AAG64297;
 XX
 DT 21-SEP-2001 (first entry)
 XX
 DE Human GTP-binding protein-coupled receptor GPRV39.
 XX


```

OY 236 KRDESLANKAVANVNIHPRSKSVTKMLFVLVFAICMPHFVDRLEFSEFVEEMTESLAA 295
    1:|||||: 11 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 244 kdkksleadeqnanlqrpcksvknmlfvlylvfalcwafphdrlffsfveeslaa 303
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY 296 VFNLHVSVGVFEYLSAVNPITLYNLSRRFAAFRRNVSPTCMKCHPRRPOGPPAQKI 355
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 304 vfolvhvsgvlfylssavnpilnyllsrrfqaafgnvlsfkhqwsghdpqlppaqrn 363
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY 356 IFLTECHLVELTEDAGPQPFPGOSSIHNTNLTTA 388
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 364 ifltechfveltedlpgqfpcqssmshnppta 396
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 10
AAB63366
ID AAB63366 standard; Protein; 415 AA.
XX
AC AAB63366;
XX
DE 15-OCT-2001 (first entry)
XX
DE Amino acid sequence of a human TGR-1 protein.
XX
KM TGR-1; neuromedin U; hypertension; stress disease.
XX
OS Homo sapiens.
XX
PN WO200157524-A1.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001: 2001WO-JP00746.
XX
PR 04-FEB-2000: 2000JP-0032773.
XX
PR 24-FEB-2000: 2000JP-0052252.
XX
PR 30-MAR-2000: 2000JP-0097896.
XX
PR 19-JUN-2000: 2000JP-0187536.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PI Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;
PI Okubo S;
XX
DR WPI: 2001-488917/53.
XX
DR N-PSDB: AAH43075.
XX
PT Identifying predicted or actual structures of two or more members of a
PT chemical or physical library by mass spectrometry comprising
PT correlating molecular mass measurements of two or more members with a
PT shared chemical history -
XX
PS Disclosure; Page 88; 95pp; Japanese.
XX
CC The present sequence represents a human TGR-1 protein. The specification
CC describes a method of screening a compound, which is capable of binding
CC properties of neuromedin U to TGR-1. The method is useful for screening
CC preventatives and remedies for hypertension, stress diseases, etc..
CC TGR-1 antagonists are also useful for treating the same diseases.
XX
SQ Sequence 415 AA:

```

```

Query Match          79.2%; Score 1644.5; DB 22; Length 415;
Best Local Similarity 79.1%; Pred. No. 5.4e-178;
Matches 311; Conservative 35; Mismatches 42; Indels 5; Gaps 1;

```

```

OY 1 MCKLENASWIIH-----DPLMKYLNSTEEYLAHLCGPKRSDLSVSAVALIFLVGMGN 55
    1:|||||: 11 ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4 meklgnaswlygqkledpfqkhnsteeylaficgprshfflpvsvvypflfvvgvign 63
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY 56 LVLWCWYVHQRHLKPTNTYVLSLAVSDLVLLGMPLEIYEMMMHNPFLFGVGCYFXT 115
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 64 vlvcwvlylqhqmkpptyllslavsdllvlllmplevemwvnyplflgpgvcyfkf 123
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

OY 116 ALFEVCFASILSVTTVSERYVAIVHPPRAKLESTRRARLRILSLVSGSVFSLPNTS 175
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 124 alfevcfasisiltcvseryvalhnpfrakqstrralrlglvwgfsvlyfsipnts 183
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY 176 HQIKFQHPNNGSSVBSATCTVTKPMWVYNIQTSTFLYILPMTLISVLYIMGLRL 235
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 184 lhgikfhyfpgngslvpgsatctvckpmylnfllyqtsflfyllpmtvlsvlylmalrl 243
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY 236 KRDESLANKAVANVNIHPRSKSVTKMLFVLVFAICMPHFVDRLEFSEFVEEMTESLAA 295
    1:|||||: 11 ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 244 kdkksleadeqnanlqrpcksvknmlfvlylvfalcwafphdrlffsfveeslaa 303
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY 296 VFNLHVSVGVFEYLSAVNPITLYNLSRRFAAFRRNVSPTCMKCHPRRPOGPPAQKI 355
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 304 vfolvhvsgvlfylssavnpilnyllsrrfqaafgnvlsfkhqwsghdpqlppaqrn 363
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
OY 356 IFLTECHLVELTEDAGPQPFPGOSSIHNTNLTTA 388
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 364 ifltechfveltedlpgqfpcqssvshnppta 396
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

RESULT 11
AAB67805
ID AAB67805 standard; Protein; 412 AA.
XX
AC AAB67805;
XX
DE 29-JUN-2001 (first entry)
XX
DE Amino acid sequence of G-protein coupled receptor IGS4B short version.
XX
KM Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
KM nervous system disorder; psychiatric disorder; Parkinson's disease;
KM episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KM cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KM motility disorder; myocardial infarction; hypertension; dyslipidemia;
KM gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KM inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KM gynecological disorder.
XX
OS Homo sapiens.
XX
PN WO200125269-A2.
XX
PD 12-APR-2001.
XX
PF 25-SEP-2000: 2000WO-EP09584.
XX
PR 24-SEP-1999: 99EP-0203140.
XX
PR 24-SEP-1999: 99NL-1013140.
XX
PR 28-JUL-2000: 2000EP-0202683.
XX
PR 31-JUL-2000: 2000US-0222047.
XX
PA (SOLV ) SOLVAY PHARM BV.
XX
PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
PI WPI: 2001-273568/28.
XX
DR N-PSDB: AAF80325.
XX
PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers -
XX
PS Claim 18; Page 89-90; 102pp; English.
XX
CC The present sequence represents the short version of a human G-protein
CC coupled receptor designated IGS4B. IGS4 exists in two polymorphic forms,
CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
CC for preventing, ameliorating or correcting dysfunctions or diseases.
CC These diseases include peripheral nervous system, psychiatric and central

```

nervous system disorders (e.g. schizophrenia, episodic paroxysmal anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or stroke), cardiovascular diseases (e.g. heart failure, angina pectoris, myocardial infarction or hypertension), dyslipidemias, obesity, emesis, gastrointestinal disorders (e.g. inflammatory bowel disease or motility disorders), osteoporosis, inflammations (e.g. bacterial, fungal, protozoan or viral), pain, cancers, immune disorders, allergies, sepsis or gynecological disorders. Agonists or antagonists of IG54 are effective with regard to disorders of the nervous system, including the central and peripheral nervous systems, disorders of the gastrointestinal system, cardiovascular system, skeletal muscle, thyroid, lung or genitourinary system, or immunological disease. The IG54 polynucleotides are useful as diagnostic reagents for detecting under-expression, overexpression or altered expression of IG54.

SQ Sequence 412 AA;

Query Match	79.0%	Score 1640.5	DB 22	Length 412
Best Local Similarity	79.1%	Pred. No. 1.5e-177		
Matches 311; Conservative	35;	Mismatches 42;	Indels 5;	Gaps 1.

```

Oy 1 MGLEMSWTH-----PMLKMYLSTREBYLAHLCGRSDLSLPVSAAYALIFLVGMGN 55
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1 meklqnsawlyqgkledpfcgkhlnsteeylaifcigprstnfiflpvavvyplfvvgvgn 60
Oy 56 LLVGVAVIRHQTLTKTPNNYULFSLAVSDLLVLLGMLPELEIEMAHNVPFLFGEVGCYFMT 115
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 61 vlvctvlqibqamktprcnnylfslavsdllvlllgnmpleyemrnyplfifgvygyfkt 120
Oy 116 ALFETVCASILSVTVTSVEREYVAIVHPRAKLESTRRRALRLSLVMSFSVVSPLMPS 175
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 121 alfevctfasllstictvseveyallhpfeklqstrrralrlglgawgsvfslpnts 180
Oy 176 IHGKFOHFPMRGSSVPSASACTYVKKPMWVNNLLIQARSEFLYLPMTLSVLTYLGLRL 235
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 181 lhgktlfhyfmgaslvpsaactctvklpmwlynfllqvstflfyllpmctvslvlylnalrl 240
Oy 236 KRDESLKANKAAVNIHRPSSRKSVYKMLFVVLVFAICTWPFPHVDRLFEFSVEEMTESLAA 295
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 241 kkdksleedegnalqprcrtsvkmwlfvlylvafctwapfhldrlfifsvveetstslaa 300
Oy 296 VFNLIHVWSGVFFYLSASAVNPDIYNLLSRFRRAFRNVSPTCKWCHPRHHPQGPAPQAKI 355
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 301 vfnlvhwvsgylyfjssavnpdlynlslstrfgaifgnvissfhkqwnsqndpqldpqrqtn 360
Oy 356 IFLTRCHLVLEITEDAGPQFPQOSSIHNTNLTTA 388
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 361 iftechchevltedipqfclqcsavhnshtpta 393

```

RESULT	12
AAB67804	
ID	AAB67804 standard; Protein; 415 AA.

AC	AAB67804;
XX	
DT	29-JUN-2001 (first entry)

DE	Amino acid sequence of a G-protein coupled receptor IGS4B long version.
XX	
KW	Human: G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
KW	nervous system disorder; psychiatric disorder; Parkinson's disease;
KW	epileptic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KW	cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW	motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW	gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW	inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW	gynecological disorder.

OS	Homo sapiens.
XX	
PN	W0200125269-A2.

XX	12-APR-2001.
PD	
XX	
PF	25-SEP-2000; 2000WO-EP09584.
XX	
PR	24-SEP-1999; 99EP-0203140.
PR	24-SEP-1999; 99NL-1013140.
PR	28-JUL-2000; 2000EP-0202683.
PR	31-JUL-2000; 2000US-0222047.
XX	
PA	(SOLV) SOLVAY PHARM BV.
XX	
PI	Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
XX	WPI; 2001-273566/28.
DR	N-PSDB; AAF80324.

PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers -

PS Claim 19; Page 85-86; 102pp; English.

The pre-m sequence represents the long version of a human G-protein coupled receptor designated IG54B. IG54 exists in two polymorphic forms, IG54A and IG54B. The IG54 receptors and IG54 polynucleotides are useful for preventing, ameliorating or correcting dysfunction or diseases. These diseases include peripheral nervous system, psychiatric and central nervous system disorders (e.g. schizophrenia, episodic paroxysmal anxiety, phobia, Parkinson's disease, migraine, epilepsy, hulaula or stroke), cardiovascular diseases (e.g. heart failure, angina pectoris, myocardial infarction or hypertension), dyslipidemias, obesity, emesis, gastrointestinal disorders (e.g. inflammatory bowel disease or motility disorders), osteoporosis, inflammations, infections (e.g. bacterial, fungal, protozoan or viral), pain, cancers, immune disorders, allergies, sepsis or gynecological disorders. Agonists or antagonists of IG54 are effective with regard to disorders of the nervous system, including the central and peripheral nervous systems, disorders of the gastrointestinal system, cardiovascular system, skeletal muscle, thyroid, lung or genitourinary system, or immunological disease. The IG54 polynucleotides are useful as diagnostic reagents for detecting under-expression, overexpression or altered expression of IG54.

50 Sequence 415 AA;

Query Match	79.08	Score 1640.5	DB 22	Length 415
Best Local Similarity	79.18	Pred. No. 1.5e-177		
Matches 311; Conservative	35	Mismatches 42	Indels 5	Gaps 1

```

0Y 1 MGRLENA5W1H-----DPLMKYLNSTERYLAHLCGPKRSDSLPVSAYALIEFLVGMGN 55
Db 4 mekiqnaaw1yqkledpfqghlnsteeya1aficlogprtshtflpvsavvyulflvgy1gn 63
0Y LLVCWAVIRHDTLKPTUYVLESLAVSODLVLGMPLEIEMHNPPLFEGPGCYFKT 115
Db 64 vlvclv1qhqamktrpny1fls1avsdlv11qmp1evemrny1pflf9pvcy1kc 122
0Y 116 ALFETVCASILSVTVTSVERVAVIWHFRAKLESTRRLARIISLVMSFSVFSLENTS 175
Db 124 alfe1vc1as1l1st1tvs1ev1y1alp1frak1q1s1rr1r1al1g1v1w1g1fs1l1p1nts 183
0Y 176 IHG1KFOHFPNGSSVGSATCTVTKPMWVYVLIIQATSELYILPMLISVLYLYLMGLRL 233
Db 184 ihg1kf1h1fpng1s1vgs1atc1tvl1kpmw1vyl1i1q1at1se1ly1il1p1ml1is1v1ly1ly1ml1rl 243
0Y KRDESLKNAKVAVNHRHSRKSVTKMLFVLVLAICMTPRHYRDLFEESPFEEMTESLAA 295
Db 244 kkdks1e1edeg1na1q1rc1r1sv1kml1fv1lv1a1lcw1ar1h1dr1l1fst1v1ee1tes1aa 303
0Y 296 VFNLIHVAVSGVFFYLSAANB1IYNL1LSRRRAAFRNVSPTCKWCPRIHNPQCPPAOKI 355

```


DT
XX
DT

PR 29-SEP-1999; 99US-0156633.
 PR 29-SEP-1999; 99US-0156634.
 PR 29-SEP-1999; 99US-0156653.
 PR 01-OCT-1999; 99US-0157280.
 PR 01-OCT-1999; 99US-0157281.
 PR 01-OCT-1999; 99US-0157282.
 PR 01-OCT-1999; 99US-0157293.
 PR 01-OCT-1999; 99US-0157294.
 PR 12-OCT-1999; 99US-0416760.
 PR 12-OCT-1999; 99US-0417044.
 XX
 PA (AREN-) ARENA PHARM INC.
 XX
 PI Chen R, Dang HT, Llaw CM, Lin I;
 XX
 DR WPI: 2000-400068/34.
 DR N-PSDB: AAD01123.
 XX
 PT Novel human orphan G protein-coupled receptors and the encoding cDNAs
 for use in the identification of G protein-coupled receptor agonists -
 XX
 PS Claim 22: Page 57-59; 102pp: English.
 XX
 CC The present amino acid sequence is the hrup6, an endogenous human
 CC orphan G protein-coupled receptor (GPCR). The full length hrup6 cDNA
 CC was cloned by RT-PCR using human thymus cDNA as template.
 CC The orphan GPCR of the invention, like all GPCRs has seven transmembrane
 CC alpha helices with an extracellular N-terminus and an intracellular
 CC C-terminus. However, no endogenous ligands have yet been identified for
 CC the proteins of the invention. The orphan GPCRs may be used in the
 CC identification of their endogenous ligands, and to screen potential GPCR
 CC agonists and antagonists for use as pharmaceutical agents. The proteins
 CC may also be used in the study of GPCR-mediated signalling cascades, and
 CC to elucidate their precise role in normal and diseased human conditions.
 CC Nucleic acid encoding human orphan GPCRs may be used for tissue
 CC localisation expression analysis to provide information about their
 CC function in healthy and pathological states.
 CC
 XX
 SO Sequence 415 AA;
 Query Match 100.0%; Score 2185; DB 21; Length 415;
 Best Local Similarity 100.0%; Pred. No. 9, 4e-246;
 Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 MSGMERKQNSMTYQOKLEDPFOKHLNSTEETYLAFGLGPRSHPELVSVYVPIFYGV 60
 Db 1 msgmeklqnaswlyqkledpfgkhlnsteeylaficgprshfiprvsvyvpilfvgy 60
 Oy 61 IGVNLVCLVILQHQAMKPTNYLYLSLAVSDLVLLGMPLEYEMWRNYPFLGPGVGY 120
 Db 61 Igvnlvclvllqhqamkptnylyfslavsdllvllgmpleyemwrnypflgpgvcy 120
 Oy 121 FKTALETFVGFASLTSTYTVSVRYVALILHFRKLOSTRRALRIIGIYMGSVLESLP 180
 Db 121 fktaletfvtfasltstlysvryvalilhfaklqstrrralriigiyvsgvlelslp 180
 Oy 181 NTSIHGKIFHYFPGSLVPSACATVIRKPMWYFIQVTSFLYLLPMTVISLYLMA 240
 Db 181 ntsihgkifhyfpgslvpsacatvirkpmwlyfiqvtsfllyllpmtvislylylma 240
 Oy 241 LRLKKDKSLADEGCMANIQRPCRSVNMKLFVLVLAICWAPPHIDRLEFSFEVEMSES 300
 Db 241 lrlkkdksladeegmaniqrpcrsvnmklfvlvlaicwapphidrlffsfveemes 300
 Oy 301 LAAVFNLYVHYVSGVFFYLSSAVNPITYNLLSRPQAAFOVVISFHKQMSQHPOLPPA 360
 Db 301 laavfnlyvhyvsgvffylssavnpitynllsrpqaafovvisfhkqmsqhpdpipa 360
 Oy 361 ORNFIETECHFEVLTEDIGPOPCQSSMNSHLPALSSQMSQSTVNTQSHFNKT 415
 Db 361 qrnfietchfevltedigpocqssmshlpalsseqmsqstvnqsfhntk 415

RESULT 2
 AAB02830
 ID AAB02830 standard; Protein; 415 AA.
 XX
 AC AAB02830;
 XX
 DT 22-AUG-2000 (first entry)
 XX
 DE Human G protein coupled receptor hrup6 protein SEQ ID NO:12.
 XX
 KW Human; G protein coupled receptor; GPCR; transmembrane receptor;
 KW identification; agonist; screening; therapeutic; pharmaceutical;
 KW mutant.
 XX
 OS Homo sapiens.
 XX
 PN WO200022131-A2.
 XX
 PD 20-APR-2000.
 XX
 PF 13-OCT-1999; 99MO-US24065.
 XX
 PR 13-OCT-1998; 98US-0170496.
 PR 12-NOV-1998; 98US-0108029.
 PR 20-NOV-1998; 98US-0109213.
 PR 27-NOV-1998; 98US-0110060.
 PR 16-FEB-1999; 99US-0120416.
 PR 26-FEB-1999; 99US-0121852.
 PR 12-MAR-1999; 99US-0123944.
 PR 12-MAR-1999; 99US-0123945.
 PR 12-MAR-1999; 99US-0123946.
 PR 12-MAR-1999; 99US-0123948.
 PR 12-MAR-1999; 99US-0123949.
 PR 12-MAR-1999; 99US-0123951.
 PR 28-MAY-1999; 99US-0136436.
 PR 28-MAY-1999; 99US-0136437.
 PR 28-MAY-1999; 99US-0136439.
 PR 28-MAY-1999; 99US-0137127.
 PR 28-MAY-1999; 99US-0137131.
 PR 28-MAY-1999; 99US-0137567.
 PR 30-JUN-1999; 99US-0141448.
 PR 27-AUG-1999; 99US-0151114.
 PR 03-SEP-1999; 99US-0152524.
 PR 29-SEP-1999; 99US-0156633.
 PR 29-SEP-1999; 99US-0156655.
 PR 29-SEP-1999; 99US-0156653.
 PR 29-SEP-1999; 99US-0156634.
 XX
 PA (AREN-) ARENA PHARM INC.
 XX
 PI Behan DP, Lehmann-Brulnasa K, Chalmers DT, Chen R, Dang HT;
 PI Gore M, Llaw CM, Lin I, Lowitz K, White C;
 XX
 DR WPI: 2000-317986/27.
 DR N-PSDB: AAA46022.
 XX
 PT Non-endogenous, human G protein-coupled receptors for screening
 receptor, inverse or partial agonists useful as therapeutic agents -
 XX
 PS Example 1: Page 86-88; 187pp: English.
 XX
 CC The present invention describes transmembrane receptors, preferably
 CC human G protein coupled receptors (GPCR), for which the endogenous
 CC ligand is unknown (orphan GPCR receptors). More specifically the present
 CC invention relates to non-endogenous, constitutively activated versions
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for
 CC the direct identification of candidate compounds as receptors agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents.
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 CC the exemplification of the present invention.
 XX
 SO Sequence 415 AA;

Query Match 100.0%; Score 2185; DB 21; Length 415;
 Best Local Similarity 100.0%; Pred. No. 9,4e-246;
 Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MSGMEKLNQASWYIYQKLEDPQKHUNSTEELAFLCGPRSHFELPVSVVYVPPIFVGV 60
 1 msqmeklqnaswlyyqkledpqkhlnsteeylaficgprtrshfllpvsvvypplfvgy 60

61 IGNVLCVLILOHQAQKPTNTNYLFSLAVSDLVLLGMPLEVEYEMRNYPFLGPGCY 120
 61 ignvlcvlilqhgaqkptntnylfslavsdllvllgmpleyemwrnypflfgpvcy 120

121 FKTALFETVCFASTLITTVSVERVVALHPFRAKLQSTRRAIRILGIYWGFSVLSLP 180
 121 fktalfetvcfaslilstvsvrvvalhpfraqlqstrralrlilgiywgfsvlslp 180

181 NTSIGIKFHFPPNGSLVPGSATCTVIKPMWIVNFIIQVTSFLFYLLPMTVISVLYYLA 240
 181 ntsigikfhyfppngslvpgsatctvikpmwivnfiiqvtsflfyllpmvvisvlyylma 240

241 LRLKDKSLDEGNANIQRCRKSVNKMFLVLYVFAICWAPPHIDRLFFSFEWSES 300
 241 lrlkkdksldegnanlqrcrksvnkmflvlylvfaicwaphidrlffsfeewses 300

301 LAAVFNLVHVSGVFYLLSSAVNPITYNLSRRFOAFQNVISFHKQWHSQHPOLPPA 360
 301 laavfnlvhvsgvfyllssavnpitynlsrrfoafqnvissfhkqwsqhpdpplpa 360

361 ORNIFLTECHVELTEDIGPOFQSSMHNHSLPTALSSQMSRTNYQSDFHFKT 415
 361 qrnifltechveltedigpofqssmhnshlptalsseqmstrnygsfhfkt 415

RESULT 3
 AAG63353 standard; Protein; 415 AA.
 ID AAG63353
 AC AAG63353:
 XX
 DT 15-OCT-2001 (first entry)
 XX
 DE Amino acid sequence of a human TGR-1 protein.
 XX
 KW TGR-1; neuromedin U; hypertension; stress disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157524-A1.
 XX
 PD 09-AUG-2001.
 XX
 PF 02-FEB-2001; 2001WO-JP00746.
 XX
 PR 04-FEB-2000; 2000JP-0032773.
 XX
 PR 24-FEB-2000; 2000JP-0052252.
 XX
 PR 30-MAR-2000; 2000JP-0097896.
 XX
 PR 19-JUN-2000; 2000JP-0187536.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H.
 XX
 PI Okubo S;
 XX
 DR WPI; 2001-488917/53.
 XX
 DR N-PSDB; AAH43072.
 XX
 PT Identifying predicted or actual structures of two or more members of a
 XX chemical or physical library by mass spectrometry compising
 XX correlating molecular mass measurements of two or more members with a
 XX shared chemical history -
 PS Claim 1; Page 77-79; 95pp; Japanese.

XX
 CC The present sequence represents a human TGR-1 protein. The specification
 CC describes a method of screening a compound, which is capable of binding
 CC properties of neuromedin U to TGR-1. The method is useful for screening
 CC preventatives and remedies for hypertension, stress diseases, etc..
 CC TGR-1 antagonists are also useful for treating the same diseases.
 XX
 SQ Sequence 415 AA:

Query Match 100.0%; Score 2185; DB 22; Length 415;
 Best Local Similarity 100.0%; Pred. No. 9,4e-246;
 Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MSGMEKLNQASWYIYQKLEDPQKHUNSTEELAFLCGPRSHFELPVSVVYVPPIFVGV 60
 1 msqmeklqnaswlyyqkledpqkhlnsteeylaficgprtrshfllpvsvvypplfvgy 60

61 IGNVLCVLILOHQAQKPTNTNYLFSLAVSDLVLLGMPLEVEYEMRNYPFLGPGCY 120
 61 ignvlcvlilqhgaqkptntnylfslavsdllvllgmpleyemwrnypflfgpvcy 120

121 FKTALFETVCFASTLITTVSVERVVALHPFRAKLQSTRRAIRILGIYWGFSVLSLP 180
 121 fktalfetvcfaslilstvsvrvvalhpfraqlqstrralrlilgiywgfsvlslp 180

181 NTSIGIKFHFPPNGSLVPGSATCTVIKPMWIVNFIIQVTSFLFYLLPMTVISVLYYLA 240
 181 ntsigikfhyfppngslvpgsatctvikpmwivnfiiqvtsflfyllpmvvisvlyylma 240

241 LRLKDKSLDEGNANIQRCRKSVNKMFLVLYVFAICWAPPHIDRLFFSFEWSES 300
 241 lrlkkdksldegnanlqrcrksvnkmflvlylvfaicwaphidrlffsfeewses 300

301 LAAVFNLVHVSGVFYLLSSAVNPITYNLSRRFOAFQNVISFHKQWHSQHPOLPPA 360
 301 laavfnlvhvsgvfyllssavnpitynlsrrfoafqnvissfhkqwsqhpdpplpa 360

361 ORNIFLTECHVELTEDIGPOFQSSMHNHSLPTALSSQMSRTNYQSDFHFKT 415
 361 qrnifltechveltedigpofqssmhnshlptalsseqmstrnygsfhfkt 415

RESULT 4
 AAG64297
 ID AAG64297 standard; Protein; 415 AA.
 XX
 AC AAG64297:
 XX
 DT 21-SEP-2001 (first entry)
 XX
 DE Human GTP-binding protein-coupled receptor GPR39.
 XX
 KW GTP-binding protein-coupled receptor; neuroprotective; immunomodulatory;
 KW muscular; urinary; circulatory; anorectic; human; guanosine triphosphate;
 XX G-protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200148189-A1.
 XX
 PD 05-JUL-2001.
 XX
 PF 28-DEC-2000; 2000WO-JP09409.
 XX
 PR 28-DEC-1999; 99JP-0375152.
 XX
 PR 31-MAR-2000; 2000JP-0101339.
 XX
 PR 23-MAY-2000; 2000JP-0155978.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Matsunoto S, Oda T, Satou Y, Morikawa N, Yoshida K, Suwa M;
 PI Sugiyama T;

The invention relates to human G-protein coupled receptors, SNORF52 and SNORF72 and their corresponding cDNA molecules. SNORF52 and SNORF72 receptors are specific for neuromedin (NMU) neuropeptides, hence they are also known as NMU receptors. The agonist and antagonist of NMU receptors are useful for treating an abnormality in a subject that is alleviated by decreasing or increasing the activity of NMU receptor. The NMU receptors serves as a valuable tool for designing drugs which are useful for treating various pathophysiological conditions such as inflammation, arthritis, autoimmune diseases, transplant rejection, graft vs host disease, bacterial, fungal, protozoan and viral infections, septicemia, AIDS, pain, psychotic and neurological disorders, including anxiety, depression, schizophrenia, dementia, mental retardation, memory loss, epilepsy, neuromotor disorders including obesity, bulimia, diabetes, eating/body weight disorders including hypertension, vascular and cardiovascular disorders, nausea, hypertension, hypotension, vascular disorders, circadian disorders, ischemia, stroke, cancers, sexual disorders, osteoporosis, benign prostatic hypertrophy, gastrointestinal disorders, nasal congestion, dermatological disorders such as psoriasis, allergies, Parkinson's disease, Alzheimer's disease, acute heart failure, angina disorders, delirium and dyskinesias such as Huntington's disease. They can also be used to regulate steroid hormone disorders, endocrine disorders, memory disorders, electrolyte balance disorders, metabolic disorders, behavioural disorders, somatosensory disorders, metabolic disorders, behavioural disorders, drug addiction, migraine, Addison's disease, Cushing's disease, prevent miscarriage, induce labour or to treat dysmenorrhoea. The present sequence is human G-protein coupled receptor, SNORF72.

SQ Sequence 415 AA:

Query Match 100.0%: Score 2185; DB 22; Length 415;
Best Local Similarity 100.0%: Pred. No. 9.4e-246;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSGMEKIQNASWYQOKLEDPFOKHLNSTEEXLAFLCGPRSHFFLPVSVVVPPIFVGV 60
DB 1 msgmeklqnaswlyqkledpfqkhnstceeylaficgprshfflpvsvvvpplfvvgv 60
OY 61 IGVNLVCLVILQOAMKPTPTNYLFLSLAVSDLLVLLGMPLEVEEMRWYPLFGPVGCY 120
DB 61 lgvnlvclvllqhqamkptptnyllfslavsdllvlllgmplevemwrnyplfgpvgy 120
OY 121 FKTALEFETWCFSASITTVSVSEYRYVALIHPRAKIQSTRRALRLGLVWGFVSFLSP 180
DB 121 fktalefetcfasilsttvsvseeryvalihpraklqstrralrlglvlgfvsflsp 180
OY 181 NTSIHGIRKHPNPGSLVPGSATCTVIRPMWYNTFIQVTSFLFYLPMPTVISLYLMA 240
DB 181 ntsihgirkhpnpgslvpgsatctvirpmyntfiiqvtsflfyllmptvislylma 240
OY 241 LRLKDKSLADEGNANIORPCRKSVMKMLFVLVLFALCMAPFHIDRLFSFVEEWSSES 300
DB 241 lrlkdksladeagnaniorpcrkssvmkmlfvlvlfalcmaphidrlfssfveewses 300
OY 301 LAAVFLVHVSGVFYSSAVNPITVNLSSRFQOAFQNVISSEFKOHNSQHDPLPPA 360
DB 301 laavflvhvsgvfylssavnpitvnlssrfqoafqnvissfkhohnsqhdplppa 360
OY 361 QRNIFLTECHFEVLETDIGPOPCOSSMHNHSLPTALSEEQMSRTVYQSFHFNKT 415
DB 361 qrnifltechfevletdigpofpcossmhnhslptalseeqmsrtvnyqsfhfnkt 415

RESULT 6
AAB67802
ID AAB67802 standard; Protein: 415 AA.

AC AAB67802;
DT 29-JUN-2001 (first entry)
DE Amino acid sequence of G-protein coupled receptor IGSA4 long version.
KW Human; G-protein coupled receptor; IGSA4; IGSA4A; IGSA4B; schizophrenia;
KW nervous system disorder; psychiatric disorder; Parkinson's disease;
KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW gynecological disorder.
OS Homo sapiens.
PN MO200125269-A2.
PD 12-APR-2001.
PF 25-SEP-2000; 2000MO-EP09584.
XX 24-SEP-1999; 99EP-0203140.
PR 24-SEP-1999; 99NL-1013140.
PR 28-JUL-2000; 2000EP-0202683.
PR 31-JUL-2000; 2000US-0222047.
XX (SOLV) SOLVAY PHARM BV.
XX Deleerenlijder W, Berger C, Loeken C, Nys G, Venema J;
XX DR WPI: 2001-273568/28.

DR N-PSDB: AAF80322.

PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers -

PS Claim 18; Page 77-79; 102pp: English.

CC The present sequence represents the long version of a human G-protein
CC coupled receptor designated IGSA4. IGSA4 exists in two polymorphic forms,
CC IGSA4 and IGSA4B. The IGSA4 receptors and IGSA4 polynucleotides are useful
CC for preventing, ameliorating or correcting dysfunctions or diseases.
CC These diseases include peripheral nervous system, psychiatric and central
CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGSA4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGSA4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGSA4.

SQ Sequence 415 AA:

Query Match 100.0%: Score 2185; DB 22; Length 415;
Best Local Similarity 100.0%: Pred. No. 9.4e-246;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSGMEKIQNASWYQOKLEDPFOKHLNSTEEXLAFLCGPRSHFFLPVSVVVPPIFVGV 60
DB 1 msgmeklqnaswlyqkledpfqkhnstceeylaficgprshfflpvsvvvpplfvvgv 60
OY 61 IGVNLVCLVILQOAMKPTPTNYLFLSLAVSDLLVLLGMPLEVEEMRWYPLFGPVGCY 120
DB 61 lgvnlvclvllqhqamkptptnyllfslavsdllvlllgmplevemwrnyplfgpvgy 120
OY 121 FKTALEFETWCFSASITTVSVSEYRYVALIHPRAKIQSTRRALRLGLVWGFVSFLSP 180
DB 121 fktalefetcfasilsttvsvseeryvalihpraklqstrralrlglvlgfvsflsp 180
OY 181 NTSIHGIRKHPNPGSLVPGSATCTVIRPMWYNTFIQVTSFLFYLPMPTVISLYLMA 240
DB 181 ntsihgirkhpnpgslvpgsatctvirpmyntfiiqvtsflfyllmptvislylma 240
OY 241 LRLKDKSLADEGNANIORPCRKSVMKMLFVLVLFALCMAPFHIDRLFSFVEEWSSES 300
DB 241 lrlkdksladeagnaniorpcrkssvmkmlfvlvlfalcmaphidrlfssfveewses 300
OY 301 LAAVFLVHVSGVFYSSAVNPITVNLSSRFQOAFQNVISSEFKOHNSQHDPLPPA 360
DB 301 laavflvhvsgvfylssavnpitvnlssrfqoafqnvissfkhohnsqhdplppa 360
OY 361 QRNIFLTECHFEVLETDIGPOPCOSSMHNHSLPTALSEEQMSRTVYQSFHFNKT 415
DB 361 qrnifltechfevletdigpofpcossmhnhslptalseeqmsrtvnyqsfhfnkt 415

RESULT 7
AAB67803
ID AAB67803 standard; Protein: 412 AA.

AC AAB67803;
DT 29-JUN-2001 (first entry)
XX

DE Amino acid sequence of G-protein coupled receptor IGS4A short version.
 XX
 XX Human: G-protein coupled receptor: IGS4: IGS4A: IGS4B: schizophrenia:
 KW nervous system disorder; psychiatric disorder; Parkinson's disease;
 KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
 KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
 KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
 KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
 KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
 KW gynecological disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200125269-A2.
 XX
 PD 12-APR-2001.
 XX
 PF 25-SEP-2000; 2000MO-EP09584.
 XX
 PR 24-SEP-1999; 99EP-0203140.
 XX 24-SEP-1999; 99NL-1013140.
 PR 28-JUL-2000; 2000EP-0202583.
 PR 31-JUL-2000; 2000US-0222047.
 XX
 PA (SOLV) SOLVAY PHARM BV.
 XX
 PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
 DR WPI; 2001-273568/28.
 DR N-PSDB; AAF80323.
 XX
 PT New G-protein coupled receptors and the polynucleotides encoding them,
 PT useful for preventing, ameliorating or correcting nervous system
 PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
 PT or cancers
 XX
 PS Claim 18: Page 81-82; 102pp; English.
 XX
 CC The present sequence represents the short version of a human G-protein
 CC coupled receptor designated IGS4A. IGS4 exists in two polymorphic forms,
 CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
 CC for preventing, ameliorating or correcting dysfunctions or diseases.
 CC These diseases include peripheral nervous system, psychiatric and central
 CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
 CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
 CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
 CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
 CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
 CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
 CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
 CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
 CC effective with regard to disorders of the nervous system, including the
 CC central and peripheral nervous systems, disorders of the gastrointestinal
 CC system, cardiovascular system, skeletal muscle, thyroid, lung or
 CC genitourinary system, or immunological disease. The IGS4 polynucleotides
 CC are useful as diagnostic reagents for detecting under-expression,
 CC overexpression or altered expression of IGS4.
 CC
 XX
 SQ Sequence 412 AA:

Query Match 99.3%; Score 2170; DB 22; Length 412;
 Best Local Similarity 100.0%; Pred. No. 5, 2e-244;
 Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 MEKLNASWITQOKLEDFQKHLNSTEYLAFLCGPRRSHFLPVSVVYPIFVGVIGN 63
 DB 1 MEKLNASWITQOKLEDFQKHLNSTEYLAFLCGPRRSHFLPVSVVYPIFVGVIGN 60
 OY 64 VIVGLVLTLOHQAOKTPNNYVLSAASDLVLLGLMPLVEYEMRANRPLFGPGCYEKT 123
 DB 61 VIVGLVLTLOHQAOKTPNNYVLSAASDLVLLGLMPLVEYEMRANRPLFGPGCYEKT 120

OY 124 ALFEIVCAFASLISITTVSEYVVAIILHPRAKLQSTRRALRIILIGWGSVLSFSPNTS 183
 DB 121 ALFEIVCAFASLISITTVSEYVVAIILHPRAKLQSTRRALRIILIGWGSVLSFSPNTS 180
 OY 184 IRIKIRHYPPNGSLVGSATCTYIKRMWITYNFIQVTSFLFYLIPMTVLSVLYMALAL 243
 DB 181 IRIKIRHYPPNGSLVGSATCTYIKRMWITYNFIQVTSFLFYLIPMTVLSVLYMALAL 240
 OY 244 KRDKSLEADCGNMIORPCRSKSVKMLFVLVFAICMPPHIDRLEFSEVESGLAA 303
 DB 241 KRDKSLEADCGNMIORPCRSKSVKMLFVLVFAICMPPHIDRLEFSEVESGLAA 300
 OY 304 VFNLVHVSGVFYSSAVNPIIYNLISRRQAQFQNTVSSFKQWHSQHDPLPAQGN 363
 DB 301 VFNLVHVSGVFYSSAVNPIIYNLISRRQAQFQNTVSSFKQWHSQHDPLPAQGN 360
 OY 364 IFLTECHFVBLTDIGPQPCQSSMHNHSLPTALSSQMSKRTVQSFHNKT 415
 DB 361 IFLTECHFVBLTDIGPQPCQSSMHNHSLPTALSSQMSKRTVQSFHNKT 412

RESULT 8

AA63366 standard; Protein; 415 AA.
 ID AA63366 standard; Protein; 415 AA.
 XX
 AC AA63366;
 XX
 DT 15-OCT-2001 (first entry)
 XX
 DE Amino acid sequence of a human TGR-1 protein.
 XX
 KW TGR-1; neuromedin U; hypertension; stress disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157524-A1.
 XX
 PD 09-AUG-2001.
 XX
 PF 02-FEB-2001; 2001MO-JP00746.
 XX
 PR 04-FEB-2000; 2000JP-0032773.
 PR 24-FEB-2000; 2000JP-0052252.
 PR 30-MAR-2000; 2000JP-0097896.
 PR 19-JUN-2000; 2000JP-0187536.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;
 PI Okubo S;
 DR WPI; 2001-488917/53.
 DR N-PSDB; AAH43075.
 XX
 PT Identifying predicted or actual structures of two or more members of a
 PT chemical or physical library by mass spectrometry comprising
 PT correlating molecular mass measurements of two or more members with a
 PT shared chemical history -
 XX
 PS Disclosure; Page 88; 95pp; Japanese.
 XX
 CC The present sequence represents a human TGR-1 protein. The specification
 CC describes a method of screening a compound, which is capable of binding
 CC properties of neuromedin U to TGR-1. The method is useful for screening
 CC preventatives and remedies for hypertension, stress diseases, etc..
 CC TGR-1 antagonists are also useful for treating the same diseases.
 CC
 XX
 SQ Sequence 415 AA:

Query Match 99.1%; Score 2166; DB 22; Length 415;
 Best Local Similarity 99.0%; Pred. No. 1, 5e-243;
 Matches 411; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSGMEKLNASWYIOOKLEDPEOKHLNSTEELAFICGPRSHFFLPVSVVYPIFVGV 60
DB 1 msgmeklgnaswlyyqgkledpfqkhnsteeeyaficgprshfflpvsvvypifvgy 60
QY 61 IGVNLVCLVILQHQAMKPTNTYLFSLAVSDLLVLLGMPLEVEYEMKRNYPFLFGVGCY 120
DB 61 IGVNLVCLVILQHQAMKPTNTYLFSLAVSDLLVLLGMPLEVEYEMKRNYPFLFGVGCY 120
QY 121 FRTALFETVCFASLITSTTVSVERVYAILHPFRAKIQSTRRRALRLIGYWGSSVLFSLP 180
DB 121 FRTALFETVCFASLITSTTVSVERVYAILHPFRAKIQSTRRRALRLIGYWGSSVLFSLP 180
QY 181 NTSIHGKIFKFFPNSGLVGSATCTVIRKPMWYFIQVTSFELYLLPMTVISLVLYLMA 240
DB 181 ntsihgkifkffpnsylvgsatctvirkpmwfyfiqvtsfelyllpmtvislvlylma 240
QY 241 LRLKKDKSLEADGNNANIQPCRSKSVNKMFLVLVLFALCWADPFHIDRLFSSVEWSES 300
DB 241 LRLKKDKSLEADGNNANIQPCRSKSVNKMFLVLVLFALCWADPFHIDRLFSSVEWSES 300
QY 301 LAAVFNLVHVSGVFYLSAVNPITYNLSRFOAFQVVISFFHKOMHSQHDPLPPA 360
DB 301 laavfnlvhvsgvfyllsavnpitynlsrfoafqvvisffhkomsqhdplppa 360
QY 361 ORNIFLTECHFEVLETDIGPQFPCOSSMHNHSLPTALSSQMSRTNYQSFFHFKT 415
DB 361 grnifltechfevletdigpqfpcgssvnmshlptalsseqmstrnyqsfhfknt 415
RESULT 9
AAB67804
ID AAB67804 standard; Protein; 415 AA.
AC AAB67804;
DT 29-JUN-2001 (first entry)
XX
DE Amino acid sequence of a G-protein coupled receptor IGS4B long version.
XX
KW Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
KW nervous system disorder; psychiatric disorder; Parkinson's disease;
KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW gynecological disorder.
OS Homo sapiens.
XX
PN WO200125269-A2.
PD 12-APR-2001.
XX
PE 25-SEP-2000; 2000WO-EP09584.
XX
PR 24-SEP-1999; 99EP-0203140.
PR 24-SEP-1999; 99WL-1013140.
PR 28-JUL-2000; 2000EP-0202683.
PR 31-JUL-2000; 2000US-0222047.
XX
PA (SOLV) SOLVAY PHARM BV.
PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
DR N-PSDB; AAF80324.
XX
PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers

XX
PS Claim 19; Page 85-86; 102pp; English.
XX
CC The present sequence represents the long version of a human G-protein
CC coupled receptor designated IGS4B. IGS4 exists in two polymorphic forms,
CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
CC for preventing, ameliorating or correcting dysfunctions or diseases.
CC These diseases include peripheral nervous system, psychiatric and central
CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
SQ Sequence 415 AA:
QY 1 MSGMEKLNASWYIOOKLEDPEOKHLNSTEELAFICGPRSHFFLPVSVVYPIFVGV 60
DB 1 msgmeklgnaswlyyqgkledpfqkhnsteeeyaficgprshfflpvsvvypifvgy 60
QY 61 IGVNLVCLVILQHQAMKPTNTYLFSLAVSDLLVLLGMPLEVEYEMKRNYPFLFGVGCY 120
DB 61 IGVNLVCLVILQHQAMKPTNTYLFSLAVSDLLVLLGMPLEVEYEMKRNYPFLFGVGCY 120
QY 121 FRTALFETVCFASLITSTTVSVERVYAILHPFRAKIQSTRRRALRLIGYWGSSVLFSLP 180
DB 121 frtalfetvcafslitsttvsvervyaailhpfraakiqstrrralrligywgsylvfslp 180
QY 181 NTSIHGKIFKFFPNSGLVGSATCTVIRKPMWYFIQVTSFELYLLPMTVISLVLYLMA 240
DB 181 ntsihgkifkffpnsylvgsatctvirkpmwfyfiqvtsfelyllpmtvislvlylma 240
QY 241 LRLKKDKSLEADGNNANIQPCRSKSVNKMFLVLVLFALCWADPFHIDRLFSSVEWSES 300
DB 241 LRLKKDKSLEADGNNANIQPCRSKSVNKMFLVLVLFALCWADPFHIDRLFSSVEWSES 300
QY 301 LAAVFNLVHVSGVFYLSAVNPITYNLSRFOAFQVVISFFHKOMHSQHDPLPPA 360
DB 301 laavfnlvhvsgvfyllsavnpitynlsrfoafqvvisffhkomsqhdplppa 360
QY 361 ORNIFLTECHFEVLETDIGPQFPCOSSMHNHSLPTALSSQMSRTNYQSFFHFKT 415
DB 361 grnifltechfevletdigpqfpcgssvnmshlptalsseqmstrnyqsfhfknt 415
RESULT 10
AAV52992
ID AAV52992 standard; Protein; 415 AA.
XX
AC AAV52992;
DT 21-FEB-2000 (first entry)
XX
DE Human neurotensin-like receptor.
XX
KW Human; neurotensin-like receptor; NLR; G-protein coupled receptor;
KW central nervous system; anaesthesia; analgesia.
XX
OS Homo sapiens.

XX MO9955733-A1.
 XX 04-NOV-1999.
 XX 15-APR-1999; 99WO-SE00598.
 XX 24-APR-1998; 98SE-0001455.
 XX (ASTR-) ASTRA PHARMA INC.
 XX (ASTR) ASTRA AB.
 XX Ahmed S, Cao J, O'Donnell D, Walker P;
 XX WPI: 2000-052803/04.
 XX N-PSDB: AAZ33297.
 XX Novel neurotensin-like receptor, useful for identifying agents for
 XX producing anesthesia or analgesia -
 XX Claim 1; Fig 2; 46pp; English.
 XX The present sequence represents a human G-protein coupled receptor
 XX neurotensin-like receptor (NLR). The NLR polynucleotide and protein
 XX can be used to isolate compounds that bind, (ant)agonise or alter
 XX the activity or expression of the NLR. The NLR is a G-protein coupled
 XX receptor which is expressed in the central nervous system and shares
 XX homology with human neurotensin receptor. The receptors can be used in
 XX assays to identify agents for producing anesthesia and analgesia.
 XX Sequence 415 AA:

Query Match 98.3%; Score 2148; DB 21; Length 415;
 Best Local Similarity 98.6%; Pred. No. 1.9e-241;
 Matches 409; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MSGMEKLNAMSWIYQOKLEDPFQKHLNSTEYLAFLCGPRSRHFFLVSVVYVPFVGVG 60
 DB 1 msnmeklqnaswlygqkledpfqkhlNSTEYLAFLCGPRSRHFFLVSVVYVPFVGVG 60
 QY 61 IGNVLVCVLIQHOAMKPTNYVYFSLAVSDLVLLGMPLEVEYEMRNVPFLGPGVCY 120
 DB 61 Ignvlvcvliqhgamkptnyylfslavsdllvllgmplyevemrnpflfpgvcy 120
 QY 121 FKTLFETVCASLITSTTVSEYRYVALHPEFRAKLOSTRRARLRILGIVGFSYLSLP 180
 DB 121 fktlfeTvcasLsttvsEYRYVALHPEFRAKLOSTRRARLRILgIVGfsYlsLP 180
 QY 181 NTSIHGIRKFNHPNGSLVPGSATCTVTKPMWITNFIIOVTSFLYLPMVIVSLVYLMA 240
 DB 181 ntsihgIRkfnHPngSLvpgsAtctvTKpmWITnfiIOvTsFLyLpMvIVsLVYLma 240
 QY 241 LRLKDSLEDEGNANIORCGRSVNKMFLVLLVFALICAPPHIDILFSPFEWSES 300
 DB 241 lrlkdsledeGNaniorCGRsvNkmFLvllVFALicAPPHIDilFSPFEwSES 300
 QY 301 LAAVFNLVHVVSGVFYLSAVNPILYVNLRSRFOAFQVNISSPHKQMSQHPOLPPA 360
 DB 301 laavfnlvHvvsGvFYlSAvNPiLYvNLRSRFOAFQvNIssPHKQMSQHPolPPa 360
 QY 361 ORNIFLECHVVELTEDIGPOFPCOSSMHNHSLPVALSSEOMSRNTQOSFHNKT 415
 DB 361 qrniflEchvELtedIGpOfPCOSSmHNhSLpVALsSEOMsrNTqOSfhnkt 415

RESULT 11
 AAB67805
 ID AAB67805 standard; Protein: 412 AA.
 XX AAB67805;
 XX 29-JUN-2001 (first entry)

XX Amino acid sequence of G-protein coupled receptor IGS4B short version.
 DE Human; G-protein coupled receptor; IGS4; IGS4B; schizophrenia;
 XX nervous system disorder; psychiatric disorder; Parkinson's disease;
 KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
 KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
 KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
 KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
 KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
 KW gynecological disorder.
 XX Homo sapiens.
 XX OS
 XX WO200125269-A2.
 XX 12-APR-2001.
 XX 25-SEP-2000; 2000WO-EP09584.
 XX 24-SEP-1999; 99EP-0203140.
 XX 24-SEP-1999; 99NL-1013140.
 XX 28-JUL-2000; 2000EP-0202683.
 XX 31-JUL-2000; 2000US-0222047.
 XX (SOLV) SOLVAY PHARM BV.
 XX Deleersnyder W, Berger C, Loeken C, Nys G, Venema J;
 XX WPI: 2001-273568/28.
 XX N-PSDB: AAF80325.
 XX New G-protein coupled receptors and the polynucleotides encoding them,
 XX useful for preventing, ameliorating or correcting nervous system
 XX disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
 XX or cancers -
 XX Claim 18; Page 89-90; 102pp; English.
 XX The present sequence represents the short version of a human G-protein
 XX coupled receptor designated IGS4B. IGS4 exists in two polymorphic forms,
 XX IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
 XX for preventing, ameliorating or correcting dysfunctions or diseases.
 XX These diseases include peripheral nervous system, psychiatric and central
 XX nervous system disorders (e.g. schizophrenia, episodic paroxysmal
 XX anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
 XX stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
 XX myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
 XX gastrointestinal disorders (e.g. inflammatory bowel disease or motility
 XX disorders), osteoporosis, inflammations, infections (e.g. bacterial,
 XX fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
 XX sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
 XX effective with regard to disorders of the nervous system, including the
 XX central and peripheral nervous systems, disorders of the gastrointestinal
 XX system, cardiovascular system, skeletal muscle, thyroid, lung or
 XX genitourinary system, or immunological disease. The IGS4 polynucleotides
 XX are useful as diagnostic reagents for detecting under-expression,
 XX overexpression or altered expression of IGS4.
 XX Sequence 412 AA;

Query Match 98.3%; Score 2147; DB 22; Length 412;
 Best Local Similarity 99.0%; Pred. No. 2.5e-241;
 Matches 408; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 4 MEKLNAMSWIYQOKLEDPFQKHLNSTEYLAFLCGPRSRHFFLVSVVYVPFVGVGN 63
 DB 1 meklqnaswlygqkledpfqkhlNSTEYLAFLCGPRSRHFFLVSVVYVPFVGVGN 60
 QY 64 VLVCLVLIQHOAMKPTNYVYFSLAVSDLVLLGMPLEVEYEMRNVPFLGPGVCYFKT 123
 DB 61 vlvclvliqhgamkptnyylfslavsdllvllgmplyevemrnpflfpgvcyfk 120

OY 364 IFLETCHEVELTEDIGPOFCOSSMHNLSHLP 396
 Db 356 iflechlmetledagpdpqpsinhnlta 388

RESULT 13

AA63367
 ID AAG63367 standard; Protein: 395 AA.

XX AAG63367;

DT 15-OCT-2001 (first entry)

DE Amino acid sequence of a rat TGR-1 protein.

XX TGR-1; neuromedin U; hypertension; stress disease.

OS Rattus sp.

XX WO200157524-A1.

XX 09-AUG-2001.

XX 02-FEB-2001; 2001WO-JP00746.

XX 04-FEB-2000; 2000JP-0032773.

XX 24-FEB-2000; 2000JP-0052252.

XX 30-MAR-2000; 2000JP-0097886.

XX 19-JUN-2000; 2000JP-0187536.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Hluma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;

XX Okubo S;

XX WPI; 2001-488917/53.

XX N-PSDB: AAA43078.

XX Identifying predicted or actual structures of two or more members of a

XX chemical or physical library by mass spectrometry comprising

XX correlating molecular mass measurements of two or more members with a

XX shared chemical history -

XX Claim 1; Page 89-91; 95pp; Japanese.

XX The present sequence represents a rat TGR-1 protein. The specification

XX describes a method of screening a compound, which is capable of binding

XX properties of neuromedin U to TGR-1. The method is useful for screening

XX preventatives and remedies for hypertension, stress diseases, etc..

XX TGR-1 antagonists are also useful for treating the same diseases.

XX Sequence 395 AA;

XX Query Match 75.2%; Score 1643.5; DB 22; Length 395;

XX Best Local Similarity 78.9%; Pred. No. 1,1e-182;

XX Matches 310; Conservative 37; Mismatches 41; Indels 5; Gaps 1;

OY 4 MEKQNSMWTQOKLEDPFOKHLSSTERYLAFLCGPRSHFFLEPVSVYVIFVVGIVGN 63

Db 1 mglklnsswln----dpImkylnsteeylahlcgprksdlsipvsyallflvgvngn 55

OY 64 VLVLVLIVHOAMKTPNRYVLESLAVSOLVLLGMPLEVEEMRNVPFELGPGVCYFKT 123

Db 56 lllvmvavtrhqltktpclnylflslavsdllvllgmpleyemhnygfltpgvcyfk 115

OY 124 ALFFETVCFASLITVTVSERYVALILHFRAKLQSTRRARILGIWGFVLSLPTS 183

Db 116 alffetvctaslsvtvseryvalilhfraklqstrrarilgisvsvfslpts 175

OY 184 IHGIRFHFPGSLVPGSATCTVIRPMWYFIQVTSFLFLLPMYIVSLVLYLMLRL 243

Db 176 ihgltkfqlfpgssavpsatctvckpmwvnylllqatsfllyllpmllsvlylmgltl 235

OY 244 KDKSLDEAGNANIORPCRSYKMLFVLYVFAICAPPHIDRLFFSFVEESLAA 303
 Db 236 kdesleankvaavnhlpsrksvklmflvlylfaicaphhivrlffsfveeslaa 295

OY 304 VFNLVHVVSGVFYLSAVNDIYNLSSRRFOAFQVNTSSFKQMSQHDPPAQRN 363

Db 296 vfnllhvvsgvffylsavndplynlssrrfrafvntvpsctkwchprbpqppaqk 355

OY 364 IFLETCHEVELTEDIGPOFCOSSMHNLSHLP 396

Db 356 iflechlmetledagpdpqpsinhnlta 388

RESULT 14

AA67806
 ID AAB67806 standard; Protein: 296 AA.

XX AAB67806;

XX 29-JUN-2001 (first entry)

XX Splice variant of G-protein coupled receptor IGS4A long version.

XX Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;

XX nervous system disorder; psychiatric disorder; Parkinson's disease;

XX episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;

XX cardiovascular disease; heart failure; angina pectoris; obesity; emesis;

XX motility disorder; myocardial infarction; hypertension; dyslipidemia;

XX gastrointestinal disorder; inflammatory bowel disease; osteoporosis;

XX inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;

XX gynecological disorder.

XX Homo sapiens.

XX WO200125269-A2.

XX 12-APR-2001.

XX 25-SEP-2000; 2000WO-EP09584.

XX 24-SEP-1999; 99EP-0203140.

XX 24-SEP-1999; 99NL-1013140.

XX 28-JUL-2000; 2000EP-0202683.

XX 31-JUL-2000; 2000US-0222047.

XX (SOLV) SOLVAY PHARM BV.

XX Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;

XX WPI; 2001-273568/28.

XX N-PSDB: AAB60326.

XX The present sequence represents a splice variant of the long version of a

XX human G-protein coupled receptor designated IGS4A. IGS4 exists in two

XX polymorphic forms, IGS4A and IGS4B. The IGS4 receptors and IGS4

XX polynucleotides are useful for preventing, ameliorating or correcting

XX dysfunctions or diseases. These diseases include peripheral nervous

XX system, psychiatric and central nervous system disorders

XX (e.g. schizophrenia, episodic paroxysmal anxiety, phobia, Parkinson's

XX disease, migraine, epilepsy, bulimia or stroke), cardiovascular diseases

XX (e.g. heart failure, angina pectoris, myocardial infarction or

XX hypertension), dyslipidemias, obesity, emesis, gastrointestinal

XX disorders (e.g. inflammatory bowel disease or motility disorders),

XX osteoporosis, inflammations, infections (e.g. bacterial, fungal,

XX protozoan or viral), pain, cancers, immune disorders, allergies,

CC sepsis or gynecological disorders. Agonists or antagonists of IG54 are
 CC effective with regard to disorders of the nervous system, including the
 CC central and peripheral nervous systems, disorders of the gastrointestinal
 CC system, cardiovascular system, skeletal muscle, thyroid, lung or
 CC genitourinary system, or immunological disease. The IG54 polynucleotides
 CC are useful as diagnostic reagents for detecting under-expression,
 CC overexpression or altered expression of IG54.

XX Sequence 296 AA:

Query Match 64.5%; Score 1409; DB 22; Length 296;
 Best Local Similarity 100.0%; Pred. No. 1.6e-155;
 Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGMEKLNASMTYQOKLEDPFOKHLNSTEYLAFLCGPRSHFFLPVSVVYPIFVGV 60
 DB 1 MSGMEKLNASMTYQOKLEDPFOKHLNSTEYLAFLCGPRSHFFLPVSVVYPIFVGV 60
 QY 61 IGVNVLVILQHQAMKPTNNYLFSLAVSDLLVLLGMPLEVEYEMRNYPFLFGVGCY 120
 DB 61 IGVNVLVILQHQAMKPTNNYLFSLAVSDLLVLLGMPLEVEYEMRNYPFLFGVGCY 120
 QY 121 FRTALFETVCFASISITTVSVRYVAIIHPFRAKQSTRRRRLRLIGTWGVSVP 180
 DB 121 FRTALFETVCFASISITTVSVRYVAIIHPFRAKQSTRRRRLRLIGTWGVSVP 180
 QY 181 NRSIHGIRKHYFPNGSLVPGSATCTYIKPMWYNFIQVTSFLFYLLPMTVISLVLYLMA 240
 DB 181 NRSIHGIRKHYFPNGSLVPGSATCTYIKPMWYNFIQVTSFLFYLLPMTVISLVLYLMA 240
 QY 241 LRLKDKSLEADGEGNANIQPCRKSVNKM 270
 DB 241 LRLKDKSLEADGEGNANIQPCRKSVNKM 270

RESULT 15

ID AAB67807 standard; Protein: 293 AA.

AC AAB67807;

DT 29-JUN-2001 (first entry)

DE Splice variant of G-protein coupled receptor IG54A short version.

XX Human; G-protein coupled receptor; IG54; IG54A; IG54B; schizophrenia;
 XX nervous system disorder; psychiatric disorder; Parkinson's disease;
 XX episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
 XX cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
 XX motility disorder; myocardial infarction; hypertension; dyslipidemia;
 XX gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
 XX inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
 XX gynecological disorder.

OS Homo sapiens.

PN WO200125269-A2.

PD 12-APR-2001.

PF 25-SEP-2000; 2000WO-EP09584.

XX 24-SEP-1999; 99EP-0203140.

PR 24-SEP-1999; 99NL-1013140.

PR 28-JUL-2000; 2000EP-0202683.

XX 31-JUL-2000; 2000US-0222047.

PA (SOLV) SOLVAY PHARM BV.

PI Delecrinijder W, Berger C, Loeken C, Nys G, Venema J;
 DR WPI; 2001-273568/28.

DR N-PSDB; AAF80327.
 XX New G-protein coupled receptors and the polynucleotides encoding them,
 PT useful for preventing, ameliorating or correcting nervous system
 PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
 PT or cancers

XX Example 1b; Page 96-97; 102pp; English.

XX The present sequence represents a splice variant of the short version of
 CC a human G-protein coupled receptor designated IG54A. IG54 exists in two
 CC polymorphic forms, IG54A and IG54B. The IG54 receptors and IG54
 CC polynucleotides are useful for preventing, ameliorating or correcting
 CC dysfunction or diseases. These diseases include peripheral nervous
 CC system, psychiatric and central nervous system disorders
 CC (e.g. schizophrenia, episodic paroxysmal anxiety, phobia, Parkinson's
 CC disease, migraine, epilepsy, bulimia or stroke), cardiovascular diseases
 CC (e.g. heart failure, angina pectoris, myocardial infarction or
 CC hypertension), dyslipidemias, obesity, emesis, gastrointestinal
 CC disorders (e.g. inflammatory bowel disease or motility disorders),
 CC osteoporosis, inflammations, infections (e.g. bacterial, fungal,
 CC protozoan or viral), pain, cancers, immune disorders, allergies,
 CC sepsis or gynecological disorders. Agonists or antagonists of IG54 are
 CC effective with regard to disorders of the nervous system, including the
 CC central and peripheral nervous systems, disorders of the gastrointestinal
 CC system, cardiovascular system, skeletal muscle, thyroid, lung or
 CC genitourinary system, or immunological disease. The IG54 polynucleotides
 CC are useful as diagnostic reagents for detecting under-expression,
 CC overexpression or altered expression of IG54.

XX Sequence 293 AA:

Query Match 63.8%; Score 1394; DB 22; Length 293;
 Best Local Similarity 100.0%; Pred. No. 8.8e-154;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MEKLNASMTYQOKLEDPFOKHLNSTEYLAFLCGPRSHFFLPVSVVYPIFVGVGN 63
 DB 1 MEKLNASMTYQOKLEDPFOKHLNSTEYLAFLCGPRSHFFLPVSVVYPIFVGVGN 60
 QY 64 VLVCLVILQHQAMKPTNNYLFSLAVSDLLVLLGMPLEVEYEMRNYPFLFGVGCY 123
 DB 61 VLVCLVILQHQAMKPTNNYLFSLAVSDLLVLLGMPLEVEYEMRNYPFLFGVGCY 120
 QY 124 ALFETVCFASISITTVSVRYVAIIHPFRAKQSTRRRRLRLIGTWGVSVP 183
 DB 121 ALFETVCFASISITTVSVRYVAIIHPFRAKQSTRRRRLRLIGTWGVSVP 180
 QY 184 IHGIRKHYFPNGSLVPGSATCTYIKPMWYNFIQVTSFLFYLLPMTVISLVLYLMA 243
 DB 181 IHGIRKHYFPNGSLVPGSATCTYIKPMWYNFIQVTSFLFYLLPMTVISLVLYLMA 240
 QY 244 KDKSLEADGEGNANIQPCRKSVNKM 270
 DB 241 KDKSLEADGEGNANIQPCRKSVNKM 267

Search completed: April 22, 2002, 14:13:39
 Job time: 52 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:19:41 ; Search time 17.05 seconds

(without alignments)
849.420 Million cell updates/sec

Title: US-09-609-146-25

Perfect score: 2076
Sequence: 1 MCKLENASWIDPLMKYLNLS.....GOSSIHNTLTAPACAGEVP 395

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	513	24.7	424	1 NTR1_RAT	P20789 rattus norv
2	508.5	24.5	424	1 NTR1_MOUSE	O88319 mus musculu
3	499	24.0	418	1 NTR1_HUMAN	P30989 homo sapien
4	477.5	23.0	412	1 GP38_HUMAN	O04393 homo sapien
5	465.5	22.4	364	1 GHSR_RAT	O08725 rattus norv
6	465	22.4	366	1 GHSR_PIG	O95254 sus scrofa
7	463.5	22.3	366	1 GHSR_HUMAN	O92847 homo sapien
8	437	21.1	410	1 NTR2_HUMAN	O95665 homo sapien
9	430.5	20.7	416	1 NTR2_RAT	O63384 rattus norv
10	426.5	20.5	417	1 NTR2_MOUSE	P70310 mus musculu
11	416	20.0	398	1 TRFR_HUMAN	P34981 homo sapien
12	414	19.9	398	1 TRFR_BOVIN	O46639 bos taurus
13	413.5	19.9	395	1 TRFR_CHICK	O93603 gallus galli
14	412	19.8	393	1 TRFR_MOUSE	P21761 mus musculu
15	412	19.8	398	1 TRFR_SHEEP	O28596 ovis aries
16	409	19.7	402	1 NK2R_CAVPO	O40777 cavia porce
17	403.5	19.4	412	1 TRFR_RAT	O01717 rattus norv
18	400.5	19.3	400	1 OPRM_HUMAN	P35372 rattus norv
19	399.5	19.2	384	1 NK2R_BOVIN	P03363 bos taurus
20	393.5	19.0	401	1 OPRM_BOVIN	P79350 bos taurus
21	393	18.9	398	1 NK2R_HUMAN	P21452 homo sapien
22	387.5	18.7	385	1 NK3R_MOUSE	P47937 mus musculu
23	387	18.6	398	1 OPRM_MOUSE	P42866 mus musculu
24	380.5	18.3	452	1 NK3R_RAT	P16177 rattus norv
25	380	18.3	398	1 OPRM_RAT	P33335 rattus norv
26	378.5	18.2	401	1 OPRM_PIG	O95247 sus scrofa
27	375.5	18.1	384	1 NK2R_RABIT	P79218 oryctolagus
28	373	18.0	384	1 NK2R_MOUSE	P30549 mus musculu
29	372	17.9	408	1 NK1R_RANCA	O98982 rana catesb
30	370.5	17.8	467	1 NK3R_RABIT	O97512 oryctolagus
31	368	17.7	390	1 NK2R_RAT	P16610 rattus norv
32	366.5	17.7	465	1 NK3R_HUMAN	P29371 homo sapien
33	363.5	17.5	384	1 SSR4_RAT	P30937 rattus norv

34	362.5	17.5	363	1 SSR5_RAT	P30938 rattus norv
35	362	17.4	380	1 OPRK_RAT	P34975 rattus norv
36	362	17.4	384	1 NK2R_MESAU	P51144 mesocricetu
37	362	17.4	453	1 GP39_HUMAN	O43194 homo sapien
38	361	17.4	504	1 TLRI1_DROME	P30974 drosophila
39	360.5	17.4	362	1 SSR5_MOUSE	O08858 mus musculu
40	360	17.3	380	1 OPRK_CAVPO	P41144 cavia porce
41	360	17.3	453	1 CCKR_XENLA	P70031 xenopus lae
42	359.5	17.3	380	1 OPRK_HUMAN	P41145 homo sapien
43	359	17.3	369	1 SSR2_HUMAN	P30874 homo sapien
44	359	17.3	440	1 NK4R_HUMAN	P30098 homo sapien
45	358	17.2	368	1 SSR2_BOVIN	P34993 bos taurus

ALIGNMENTS

RESULT	ID	STANDARD	PRT	424 AA.
1	NTR1_RAT			
AC	P20789:			
DT	01-FEB-1991 (rel. 17, Created)			
DT	01-FEB-1991 (rel. 17, Last sequence update)			
DT	15-JUL-1998 (rel. 36, Last annotation update)			
DE	NEUROTENSIN RECEPTOR TYPE 1 (NT-R-1) (HIGH-AFFINITY LEVOCABASTINE-SENSITIVE NEUROTENSIN RECEPTOR) (NTRH).			
GN	NTR1.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=90297956; PubMed=1694443;			
RA	Tanaka K., Masu M., Nakanishi S.;			
RT	"Structure and functional expression of the cloned rat neurotensin receptor."			
RT	Neuron 4:847-854(1990).			
RL	-1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.			
CC	HIGHEST TO TACHYKININS RECEPTORS.			
DR	PIR: JH0164; JH0164.			
DR	GCRD: GCR_0219; -			
DR	InterPro: IPR000276; GPCR_Rhodpsn.			
DR	Pfam: PF00001; 7tm_1; 1.			
DR	PRINTS: PS00237; GPCR_RHODOPSIN.			
DR	PROSITE: PS00237; G-PROTEIN_RECEP_FL_1; 1.			
DR	PROSITE: PS50262; G-PROTEIN_RECEP_FL_2; 1.			
KW	G-protein coupled receptor; Transmembrane; Glycoprotein;			
KW	Phosphorylation; Lipoprotein; Palmitate.			
FT	DOMAIN 1 64			
FT	TRANSMEM 65 87			
FT	DOMAIN 88 96			
FT	TRANSMEM 97 121			
FT	DOMAIN 122 143			
FT	TRANSMEM 144 165			
FT	DOMAIN 166 188			
FT	TRANSMEM 189 210			
FT	DOMAIN 211 235			
FT	TRANSMEM 236 260			
FT	DOMAIN 261 308			
FT	TRANSMEM 309 330			
FT	DOMAIN 331 348			
FT	TRANSMEM 349 372			
FT	DOMAIN 373 424			
FT	CAROHND 3 4			
FT	CAROHND 38 38			
FT	CAROHND 42 42			
FT	DISULFID 142 225			

FT LIPID 388 388 PALMITATE (POTENTIAL).
SQ SEQUENCE 424 AA: 47054 MW: A9C2F7EAF8D9BCD3 CRC64;
Query Match 24.7%; Score 513; DB 1; Length 424;
Best Local Similarity 31.6%; Pred. No. 2,6e-26;
Matches 119; Conservative 76; Mismatches 113; Indels 68; Gaps 13;
OY NSTEEYLALHLCGPKRSLSLP-----VSVALFLFLVGVGNLLVCMVIVRH--QT 67
DB 42 NTSESQDA---GP NSLDVNTDIYSKVLYTAIYALFLVGVGNSTVATFLARKKSLSQS 97
OY 68 LKPTNYLYESLAVSDLLVLLGMPLEIYE-MMHNPFLGPGVC---YEKTALEFYVCF 123
DB 98 LOSTVHHGLASLSDLLILLAMPVELYNFIWHHPARGADGCRGYF---LRDACTY 154
OY 124 ASLISVTYVERVVALVHFRPAKLESTRRALRILSLVSVSVSFLPMTSHIGIKFOH 183
DB 155 ATALNVALSLVERVLAICHFKATLMSRSRKTKEFISAILASALALPMLFTMGL--QN 212
OY 184 FPNQSSVPGSATCTVTKPMVNYLIIOATSEFLFYILPMTLISVLYLMGLRLKRDSELEA 243
DB 213 RSGDGHFGGLVCTPIYDITAVKVIQVNTFMSEFLPMLVLSIL-----NTVIA 261
OY 244 NKVAVINIR-----PSRKSVTK---MLFVLVLAFCW 273
DB 262 NKLTVMHVQAEOGRVCTVGHNGLEHSTFNMTIEPGVQALRHGVLAIVAVIAFVVCW 321
OY 274 TPPIHVDRLEFSFV--EEMTESLAVENLIVSVGFYLSAVNPITLYNLSRFRFAFR 331
DB 322 LPIYHRLKMFCTYSDDEMTTFLEDFYHYFYMALNALFYSSAINPILYLNVSANFRQVFL 381
OY 332 NVVSPCK-WCHPRHR 346
DB 382 STLACLCPGMRHRKK 397
RESULT 2
NTRL_MOUSE
ID NTRL_MOUSE STANDARD; PRT: 424 AA.
AC 088319:
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE NEUTROGENIN RECEPTOR TYPE 1 (NT-R-1).
GN NTSR1 OR NTSR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Snider J., Sano H., Ohta M.;
RT Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
RL - FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROGENIN. IT IS
CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHYKININS RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AB017027; BAA33013.1; -
DR MGD; MGI:97386; Ntsr.

DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; 1.
DR PROSITE: PS50262; G_PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 64
FT TRANSMEM 65 87
FT DOMAIN 88 96
FT TRANSMEM 97 121
FT DOMAIN 122 143
FT TRANSMEM 144 165
FT DOMAIN 166 188
FT TRANSMEM 189 210
FT DOMAIN 211 235
FT TRANSMEM 236 260
FT TRANSMEM 261 308
FT TRANSMEM 309 330
FT DOMAIN 331 348
FT TRANSMEM 349 372
FT DOMAIN 373 424
FT CARBOHYD 4 4
FT CARBOHYD 38 38
FT CARBOHYD 42 42
FT CARBOHYD 211 211
FT DISULFID 141 224
FT LIPID 388 388
SQ SEQUENCE 424 AA: 47216 MW: 8E9A723171A48711 CRC64;
PALMITATE (POTENTIAL).
Query Match 24.5%; Score 508.5; DB 1; Length 424;
Best Local Similarity 32.2%; Pred. No. 5,1e-26;
Matches 116; Conservative 70; Mismatches 99; Indels 75; Gaps 13;

OY 40 VSVAYALIFLVGMNLLVCMVIVRH--QTLKPTNYLYESLAVSDLLVLLGMPLEIY 96
DB 66 VTAIVYLLFLFYGVGVGNSTVATFLARKKSLSQSLSQSTVHHGLASLSDLLILLAMPVELY 125
OY 97 E-MMHNPFLGPGVC---YEKTALEFYVCFASLSTTYVSRVYVAVIYHFRPAKLESTR 152
DB 126 NFIWVHHHPMAFGAGCGGYF---LRDACTYAVALNVALSVVERVLAICHFKATLMSR 182
OY 153 RRALRILSLVSVSVSFLP-----NTSINGIKFOHPNCSVPGSATCTVTKPMV 204
DB 183 SRTRKFTSAITWALSALLAVVMTFMGLQNRASAG---QH-----PGGLVCTPIYDTAT 232
OY 205 YNLIOATSEFLYILPMTLISVLYLMGLRLKRDSELEANKVAVNIRH----- 252
DB 233 VKVVIQVNTFMSEFLPMLISIL-----NTVIANKLTVMHVQAEOGRGVCTVG 281
OY 253 -----PSRKSVTK---MLFVLVLAFCWTPPIHVDRLEFSFV--EEMTE 291
DB 282 THNSLEHSTFNMSIEGRVQALRHGVLAIVAVIAFVVCWLPYHRLKMFCTYSDQWTT 341
OY 292 SLAAVENLIVHGVGFYLSAVNPITLYNLSRFRFAFNVSPTCK-WCHPRHRPGCF 350
DB 342 FLDFDFHYFFMLNALFYSSAINPILYLNVSANFRQVFLSTLACLCPGM--RRRRKKRP 399
RESULT 3
NTRL_HUMAN
ID NTRL_HUMAN STANDARD; PRT: 418 AA.
AC P30989:
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NEUTROGENIN RECEPTOR TYPE 1 (NT-R-1) (HIGH-AFFINITY LEVOCABASTINE-
DE INSENSITIVE NEUROGENIN RECEPTOR) (NTNH).
GN NTSR1 OR NTSR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93154505; PubMed=8381365;
RA Vita N., Laurent P., Lefort S., Chalou P., Dumont X., Kaghad M.,
RA Gully D., Le Fur G., Ferrara P., Caput D.;
RT Cloning and expression of a complementary DNA encoding a high
RT affinity human neurotensin receptor."
RL FEBS Lett. 317:139-142(1993).
CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TCHYKININS RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X70070; CAA49675.1; -.
DR PIR: S29506; S29506.
DR GCRDB: GCR_0577; -.
DR GCRDB: GCR_2067; -.
DR MIM: 162651; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PROSITE: PS00237; G-PROTEIN_RECEP_FL_1; 1.
DR PROSITE: PS50262; G-PROTEIN_RECEP_FL_2; 1.
KM G-protein coupled receptor; Transmembrane; Glycoprotein;
KM Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 63
FT TRANSMEM 64 86
FT DOMAIN 87 95
FT TRANSMEM 96 120
FT DOMAIN 121 142
FT TRANSMEM 143 164
FT DOMAIN 165 187
FT TRANSMEM 188 209
FT DOMAIN 210 234
FT TRANSMEM 235 259
FT DOMAIN 260 303
FT TRANSMEM 304 325
FT DOMAIN 326 343
FT TRANSMEM 344 367
FT DOMAIN 368 418
FT CARBOHYD 4 4
FT CARBOHYD 37 37
FT CARBOHYD 41 41
FT DISULFID 141 224
FT LIPID 383 383
SQ SEQUENCE 418 AA; 46288 MW; BBBD1EEC2BE6E390 CRC64;

```

```

Query Match 24.0%; Score 499; DB 1: Length 418;
Best Local Similarity 33.0%; Pred. No. 2e-25;
Matches 116; Conservative 65; Mismatches 103; Indels 68; Gaps 12;

```

```

OY 205 YNLIQATSPFLYILPMTLISLYLMLGLRKDESLKANKAVNIHR----- 252
DB 233 VKVIVQNTFMSTFFPVAVISLV-----NTIIANKLIVNVRQAAGOCVTVGS 281
OY 253 -----PSR-----KSVTKMLFVLVLEAICMPFHDRLFESFV--EEMTESLAIV 296
DB 282 EHSFMSAIEPGHVRQALRHGVRLRAVIAVFAVCWMLPYHVRIMPCYISDEQWTPFLYDP 341
OY 297 ENLIHYVSGFFFLSSAVNPITINLSRRRAFRNVSTGK-WCHPRHRP 347
DB 342 YHVEYVNTNALFVYSTINPILNVLVANSRHFLEATLACLPVWRRRRKRP 393

```

RESULT 4

GP38_HUMAN STANDARD: PRT; 412 AA.

AC 043193;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 20-AGU-2001 (Rel. 40, Last annotation update)

DE PUTATIVE G-PROTEIN-COUPLED RECEPTOR GPR38.

GN GPR38.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98110578; PubMed=9441746;

RA McKee K.K., Tan C.P., Palaya O.C., Liu J., Feighner S.D.,

RA Hreniuk D.L., Smith R.G., Howard A.D., van der Ploeg L.H.T.;

RT Cloning and characterization of two human G-protein-coupled receptor

RT genes (GPR38 and GPR39) related to the growth hormone secretagogue

RT and neurotensin receptors."

RL Genomics 46:426-434(1997).

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN THYROID, STOMACH, AND BONE

CC MARROW.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AF034632; AAC26081.1; -.
DR GCRDB: GCR_2494; -.
DR MIM: 602885; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 2.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS00237; G-PROTEIN_RECEP_FL_1; 1.
DR PROSITE: PS50262; G-PROTEIN_RECEP_FL_2; 1.
KM G-protein coupled receptor; Transmembrane; Glycoprotein;
KM Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 35
FT TRANSMEM 36 56
FT DOMAIN 57 74
FT TRANSMEM 75 94
FT DOMAIN 95 112
FT TRANSMEM 113 134
FT DOMAIN 135 157
FT TRANSMEM 158 178
FT DOMAIN 179 246
FT TRANSMEM 247 270
FT DOMAIN 271 298
FT TRANSMEM 299 320
FT DOMAIN 321 334
FT TRANSMEM 335 358
FT DOMAIN 359 412

FT DISULFID 111 235 BY SIMILARITY.
 FT CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 412 AA; 45344 MW; C13FF6165012DEF3 CRC64;

Query Match 23.0%; Score 477.5; DB 1; Length 412;
 Best Local Similarity 30.8%; Pred. No. 4.8e-24;
 Matches 112; Conservative 81; Mismatches 114; Indels 57; Gaps 9;

QY 29 CGKRRLSLPVSVAVALFLVGVGMGLVCMVYVRHQTLPNTNYLFLSLAVSDLLVL 88
 DB 30 CSPFLPALVPLVAVCLFLFVGVGVGVVTVMLGRYDRKRTNLTLSGMAVSDDLIL 88
 QY 89 LGPMLLEYEMHNVPFLFPGVGVCFKALFETVCFASILSTVTSVRVAIVPPRAKL 148
 DB 89 LGPFLDLYRMRKRPWFGLCLRLSLYVGGCYATLTMHTALSVRYALICPLRARV 148
 QY 149 ESTRRRLRLTLVSPSVFSLPNTSHIGIKFQHPNGSSVPG-----SATCTVTK 200
 DB 149 LVTRRRVRLAIVLMAVALLSAGPFLVGV--EQDPGISVPLNGTARIASSPLASP 206
 QY 201 PMWV-----YLLIQAISFLFIYLPMLISV 226
 DB 207 PLMLSRAPPSPPSGPETAALALFSRECRSPAQLGALRYMLMWTVAFF-LPFLCLSI 265
 QY 227 LVYLMGLRLKRDLSLEANKVAVNIHRPSKSVTKMLFVLVLFVAFICWTPHYDRLFPSV 286
 DB 266 LVYLGELMLSSRRPLRGPASGREGHROT-V-LVLLVVLAFITCLPFRVGHIIYINT 324
 QY 287 EEMTESLAAVFNLIHVSGVFYLSAVNPITYNLSRRRR-AAFRNVSPTECKMCHPRH 345
 DB 325 ED--SRMVFYSQYFNIVALQLFLYLSASINPILYLSKRYRAAFKLLA-----RKS 375
 QY 346 RPOG 349
 DB 376 RPKG 379

RESULT 5

GHSR_RAT STANDARD; PRT; 364 AA.

AC 008725;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GROWTH HORMONE SECRETAGOCUE RECEPTOR TYPE 1 (GHS-R) (GH-RELEASING PEPTIDE RECEPTOR) (GHRP).
 GN GHSR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Pituitary;
 RX MEDLINE=97246555; PubMed=9092793;
 RA McKee K.K., Palyha O.C., Feighner S.D., Hreniuk D.L., Tan C., Phillips M.S., Smith R.G., der Ploeg L.H.T., Howard A.D., "Molecular analysis of rat pituitary and hypothalamic growth hormone secretagogue receptors."
 RT Molecular analysis of rat pituitary and hypothalamic growth hormone secretagogue receptors."
 RL MOL. Endocrinol. 11:415-423(1997).

RP SEQUENCE OF 1-240 FROM N.A.
 RC STRAIN-Wistar; Tissue-Pituitary;
 RX MEDLINE=98100386; PubMed=9437732;
 RA Yokota R., Sato M., Matsubara S., Ohye H., Mimi M., Murao K., Takahara J.;
 RT "Molecular cloning and gene expression of growth hormone-releasing peptide receptor in rat tissues."
 RL Peptides 19:15-20(1998).
 CC - FUNCTION: RECEPTOR FOR GROWTH HORMONE RELEASING PEPTIDES (GHRP) AS WELL AS NON-PEPTIDE, LOW MOLECULAR WEIGHT SECRETAGOGUES (E.G. L-

CC 692,429, MK-0677). THIS RECEPTOR IS COUPLED TO G-ALPHA-11
 CC PROTEIN (BY SIMILARITY).
 CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN
 CC - SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: U94321; AAC5156.1; -
 DR EMBL: AB001982; BAA2177.1; ALI_INIT.
 DR GCRDB: GCR_1383; -
 DR InterPro: IPR003905; GHS1_receptor.
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCRHOOPS.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 40
 FT TRANSMEM 41 66
 FT DOMAIN 67 72
 FT TRANSMEM 73 96
 FT DOMAIN 97 117
 FT TRANSMEM 118 139
 FT DOMAIN 140 162
 FT TRANSMEM 163 183
 FT DOMAIN 184 211
 FT TRANSMEM 212 235
 FT DOMAIN 236 263
 FT TRANSMEM 264 285
 FT DOMAIN 302 302
 FT TRANSMEM 303 326
 FT DOMAIN 327 364
 FT DISULFID 115 197
 FT CARBOHYD 13 13
 FT CARBOHYD 26 26
 SQ SEQUENCE 364 AA; 40963 MW; DCBF539BE061EEB9 CRC64;

Query Match 22.4%; Score 465.5; DB 1; Length 364;
 Best Local Similarity 34.6%; Pred. No. 2.5e-23;
 Matches 104; Conservative 65; Mismatches 107; Indels 25; Gaps 6;

QY 40 VSAVALFLVGVGMGLVCMVYVRHQTLPNTNYLFLSLAVSDLLVLGMPLEYEMW 99
 DB 45 VTATCVALLFVGVISGNLLTMVLSRFRRLRTTNLYLSMAFSDDLIFLC-MPLDLVRLM 103
 QY 100 HNPFLFPGVGCYFKTALFETVCFASILSVTVSVRYVAIVHPFRKLESTRRLRL 159
 DB 104 QYRPMNFGDLCKLQFVSSCYATVLTITLALSVERFYAICPLRAKVVTGGRVLYI 163
 QY 160 SLVVSFVSFLPMTSHIGIKFQHPNGSSVPGSATCTVTK-----PMWVYLLIT 209
 DB 164 LVITAAVAFCSAGPFLVGVGEH---NGTDPDPTNEGRATLFAVRSGLLTVMWV----- 215
 QY 210 QATSEFLFIYLPMLISVLYLMGLRLKRDLSLEANKVAVNIHRPSKSVTKMLFVLVLF 269
 DB 216 ---SVVFPLPVCLTVLTVLSLIGRKLRRGGDA--VGASLRDQNHQYKMLAVVFAF 270
 QY 270 AICHTPFRHYDRLFFS-FVEMTESLAAVFNLIHVSGVFYLSAVNPITYNLSRRFR 328
 DB 271 ILCMLPFRHYDRLFFSKSEFEGSLIAQISQYCNLVSEFLVFLYLSAINDPIYINMSKRYV 330
 QY 329 A 329
 DB 331 A 331

	CC	-1- SUBCELLULAR LOCATION:	INTRAGOL MEMBRANE PROTEIN.
	CC	-1- ALTERNATIVE PRODUCTS:	2 ISOFORMS; 1A (SHOWN HERE) AND 1B: ARE PRODUCED BY ALTERNATIVE SPLICING. ISOFORM 1B APPEARS NOT TO BIND SECRETAGOGUES.
	CC	-1- TISSUE SPECIFICITY:	PITUITARY AND HYPOTHALAMUS.
	CC	-1- SIMILARITY:	BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
	CC		-----
	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb.ch/announce/ or send an email to license@isb.slb.ch).	
	CC		-----
	DR	EMBL;	U60179; AAC50653.1; -.
	DR	EMBL;	U60181; AAC50654.1; -.
	DR	GCRDB;	GCR_1917; -.
	DR	GCRDB;	GCR_1918; -.
	DR	MIM;	601898; -.
	DR	InterPro;	IPR003905; GHSL_Rcceptor.
	DR	InterPro;	IPR000276; GPCR_Rhodopsn.
	DR	PIfam;	PF00001; 7tm1r.1.
	DR	PRINTS;	PR00337; GPCRRHODOPSN.
	DR	PROSITE;	PS00237; G_PROTEIN_RECP_F1.1; 1.
	DR	PROSITE;	PS50262; G_PROTEIN_RECP_F1.2; 1.
	KW		G-protein coupled receptor; Transmembrane; Glycoprotein; Alternative splicing.
FT	DOMAIN	1	40
FT	TRANSMEM	41	66
FT	DOMAIN	67	72
FT	TRANSMEM	73	96
FT	DOMAIN	97	117
FT	TRANSMEM	118	139
FT	DOMAIN	140	162
FT	TRANSMEM	163	183
FT	DOMAIN	184	211
FT	TRANSMEM	212	235
FT	DOMAIN	236	263
FT	TRANSMEM	264	285
FT	DOMAIN	286	302
FT	TRANSMEM	303	326
FT	DOMAIN	327	366
FT	DISEULFD	116	198
FT	CARBOHYD	13	13
FT	CARBOHYD	27	27
FT	VANSPPLC	266	289
FT	VANSPPLC	290	366
DO	SEQUENCE	366 AA;	41328 MW; D1B6z71DDA9DC06 CRC64; MISSING (IN ISOFORM 1B);

[illegible]

DB	QY	Db	RESULT	ID	NTFR2_HUMAN	STANDARD:	PRT:	410 AA.
273	328 AAFR 331	333	NTFR2_HUMAN	AC 095655:	20-AUG-2001 (Rel. 40, Created)			
	1 1 1				20-AUG-2001 (Rel. 40, Last sequence update)			
					20-AUG-2001 (Rel. 40, Last annotation update)			
					NEUROKININ RECEPTOR TYPE 2 (NT-8-2) (LEVOCABASTINE-SENSITIVE			
					NEUROKININ RECEPTOR (NTR2 RECEPTOR).			
					NTSR2.			
					OS Homo sapiens (Human).			
					Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
					Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
					NCBI_Taxid:9606;			
					[1]			
					SEQUENCE FROM N.A.			
					RC TISSUE=Brain;			
					RA MEDLINE=99066919; PubMed=9851594;			
					RA Vita N., Oury-Donat F., Chalou P., Guillemot M., Kaghad M., Bachy A.,			
					RA Thuryssen S., Garcia S., Polnot-Chazel C., Casellas P., Keane P.,			
					RA Le Fur G., Maffrand J.P., Shoubrie P., Caput D., Ferrara P.:			
					RT "Neurotensin is an antagonist of the human neurotensin NT2 receptor			
					RT expressed in Chinese hamster ovary cells.";			
					RL Eur. J. Pharmacol. 360:265-272(1998).			
					CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS			
					CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-			
					CC CALCIUM SECOND MESSENGER SYSTEM.			
					CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.			
					CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.			
					CC HIGHEST TO TACHYKININS RECEPTORS.			
					CC -----			
					CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
					CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
					CC the European Bioinformatics Institute. There are no restrictions on its			
					CC use by non-profit institutions as long as its content is in no way			
					CC modified and this statement is not removed. Usage by and for commercial			
					CC entities requires a license agreement (See http://www.isb-sib.ch/announce/			
					CC or send an email to license@isb-sib.ch).			
					CC -----			
					CC EMBL: Y10148; CAA71233.1; -.			
					DR MIT: 605538; -.			
					DR InterPro: IPR000276; GPCR_Rhodpsn.			
					DR Pfam: PF00001; 7tm_1; 1.			
					DR PRINTS: PR00237; GPCRHRHODPSN.			
					DR PROSITE: PS00237; G_PROTEIN_RECPT_FL_1; 1.			
					DR PROSITE: PS00262; G_PROTEIN_RECPT_FL_2; 1.			
					KW G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.			
					FT DOMAIN 1 32			
					FT TRANSMEM 33 55			
					FT DOMAIN 56 64			
					FT TRANSMEM 65 87			
					FT DOMAIN 88 109			
					FT TRANSMEM 110 131			
					FT DOMAIN 132 154			
					FT TRANSMEM 155 176			
					FT DOMAIN 177 217			
					FT TRANSMEM 218 237			
					FT DOMAIN 238 297			
					FT TRANSMEM 298 318			
					FT DOMAIN 319 337			
					FT TRANSMEM 338 358			
					FT DOMAIN 359 410			
					FT DISULFID 108 194			
					FT LIPID 377 377			
					SEQUENCE 410 AA; 45413 MW; 8C3AA22BE15FD66 CRC64;			

Query Match 21.1%; Score 437; DB 1; Length 410;
 Best Local Similarity 30.2%; Pred. No. 1.9e-21;
 Matches 108; Conservative 73; Mismatches 123; Indels 54; Gaps 11;

CC 44 VALIFLVGMNLLVCMVIRHOTLKT-PTNYLFSLAVSDLLVLLGPLEIYE-MMHN 101
 DB 39 VALINLGAAGNALSVHYLAKRAGRRLRHVLSLALAGLLLVGPVELYSFVWH 98
 QY 102 YPFLGPGVC---YFKTALFETVCASILSVTVSYVERVAIVHPRALESTRRALRI 158
 DB 99 YPWFVGLDCCRGYF---VHELCAVATVLSVAGLSAERCLACOPRLARSLTLPRTTRL 155
 QY 139 LSLVMSFSVFSFLPNTSHIGIKFO-HFPNGSSVPGSATCTVTKPMNVNLLIQTSTFLRY 217
 DB 136 VALSNASISGLALPMAVIMGOKHELETADDEPSPASRVTCLVSRALQVFIQVNVLSF 215
 QY 218 ILPMTLLISVL-----YLLMG-----LRLKRESL-----BA 243
 DB 216 VLPLALTLAFNLGVTSLLALCSQVPSSTPGSSPSPRLLESEGLLFIWKKTFIG 275
 QY 244 NKAVANIHPRSK-----SVTKMLEVLYLVFAICTPFRVVDLRFSEFV--EEKTESLA 296
 DB 276 GCVSLVRHKDVRRISLQSRVQLRAIVAVYICMLPYHARRLMCYVPDAMTDPLNFR 335
 QY 297 FLNLIHVSGVFYLSAVNPITLYNLSRRERAFAFRNVVSPCKMCHPRHR---PGCP 350
 DB 336 YHFFVMTVTLTFYVSASVATPLILYNAVSSFRKLFLEAVSSLCGEHHPMKRLPPKPOSP 393

RESULT 9
 NTR2_MOUSE STANDARD: PRT: 416 AA.

AC 06384:
 DT 01-NOV-1997 (rel. 35, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 20-AUG-2001 (rel. 40, Last annotation update)
 DE NEUROTENSIN RECEPTOR TYPE 2 (NT-R-2) (HIGH-AFFINITY LEVOCABASTINE-SENSITIVE NEUROTENSIN RECEPTOR).
 GN NTR2 OR NTR2.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Hypothalamus;
 RX MEDLINE=96228041; PubMed=8647296;
 RA Chalon P., Vita N., Kaghad M., Guillemont M., Bonin J.,
 RA Delpech B., Le Fur G., Ferrara P., Caput D.;
 RT "Molecular cloning of a levocabastine-sensitive neurotensin binding site."
 RT FEBS Lett. 386:91-94(1996).
 CC - FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
 CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC - TISSUE SPECIFICITY: ABUNDANT IN CORTX AND HYPOTHALAMUS, AND LOWER
 CC LEVELS SEEN IN THE HEART AND INTESTINE.
 CC - DEVELOPMENTAL STAGE: EXPRESSED MAXIMALLY IN 7-DAY-OLD BRAIN AND
 CC EXPRESSION DECREASES PROGRESSIVELY UNTIL ADULTHOOD (35-DAY-OLD
 CC BRAIN).
 CC - SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHYKININS RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).

CC 44 VALIFLVGMNLLVCMVIRHOTLKT-PTNYLFSLAVSDLLVLLGPLEIYE-MMHN 101
 DB 39 VALIFLAFGTAGNALSVHYLAKRAGRGRRLRHVLSLALSALLVSMELTNFVWSH 98
 QY 102 YPFLGPGVC---YFKTALFETVCASILSVTVSYVERVAIVHPRALESTRRALRI 158
 DB 99 YPWFVGLDCCRGYF---VRELCAVATVLSVAGLSAERCLACOPRLARSLTLPRTTRL 155
 QY 159 LSLVMSFSVFSFLPNTSHIGIKFO-HFPNGSSVPGSATCTVTKPMNVNLLIQTSTFLRY 217
 DB 156 LSLVMSVSLGLALPMAVIMGOKHEVSADDEPSPASRVTCLVSRATLQVFIQVNVLSF 215
 QY 218 ILPMTLLISVL-----YLLMG-----LRLKRESL-----BA 239
 DB 216 ALPLALTLAFNLGVTSLLALCSQVPSASQVSSIPRLLESEGLLGFITWRKTLISG 275
 QY 240 ---SLANKAVANIHPRSKSVTKMLFVLVFAICTPFRVVDLRFSEFVEE--WTESLA 294
 DB 276 VQASLVRRHKDASQIR--SLQHSQVQLRAIVAVYICMLPYHARRLMCYTPDDGWTNEL 333
 QY 295 AVENLIHVSGVFYLSAVNPITLYNLSRRERAFAFRNVVSPCKMCH---PRHRPQ 348
 DB 334 DFHYHFMVMTLTFYVSASVATPLILYNAVSSFRKLFLESIGSGHOSLVRPLQAPAE 392

RESULT 10
 NTR2_MOUSE STANDARD: PRT: 417 AA.

AC P70310:
 DT 01-NOV-1997 (rel. 35, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 20-AUG-2001 (rel. 40, Last annotation update)
 DE NEUROTENSIN RECEPTOR TYPE 2 (NT-R-2) (LOW-AFFINITY LEVOCABASTINE-SENSITIVE NEUROTENSIN RECEPTOR) (NTR1).
 GN NTR2.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

Query Match 20.7%; Score 430.5; DB 1; Length 416;
 Best Local Similarity 31.5%; Pred. No. 4.9e-21;
 Matches 113; Conservative 61; Mismatches 126; Indels 59; Gaps 11;

RC STRAIN-BALB/C; TISSUE-Brain;
 RA MEDLINE-96388216; PubMed-8795617;
 RA Mazella J., Boto J.-M., Guillemane E., Coppola T., Sarret P.,
 RA Vincent J.-P.;
 RT "Structure, functional expression, and cerebral localization of the
 RT levoocadherine-sensitive neurotensin/neuromedin N receptor from mouse
 RT brain.";
 RL J. Neurosci. 16:5613-5620(1996).
 CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEURENSEN. IT IS
 CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: EXPRESSED MAXIMALLY IN THE CEREBELLUM,
 CC HIPPOCAMPUS, PIRIFORM CORTEX AND NEOCORTX OF ADULT BRAIN.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED POORLY IN 7-DAY-OLD BRAIN.
 CC EXPRESSION INCREASES AT DAY 15 TO REACH A MAXIMAL LEVEL IN 35-DAY-
 CC OLD BRAIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHYKININS RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U51908; AAB17285.1; -
 DR MGD: MGI:108018; Nstr2.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 2.
 DR PROSITE: PS00237; G-PROTEIN_RECPT_F1.1; 1.
 DR PROSITE: PS00262; G-PROTEIN_RECPT_F1.2; 1.
 KW G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.
 FT DOMAIN 1 32
 FT TRANSMEM 33 55
 FT DOMAIN 56 64
 FT TRANSMEM 65 87
 FT DOMAIN 88 109
 FT TRANSMEM 110 131
 FT DOMAIN 132 154
 FT TRANSMEM 155 176
 FT DOMAIN 177 217
 FT TRANSMEM 218 238
 FT DOMAIN 239 298
 FT TRANSMEM 299 319
 FT DOMAIN 320 338
 FT TRANSMEM 339 359
 FT DOMAIN 360 417
 FT DISULFID 108 194
 FT LIPID 378
 SQ SEQUENCE 417 AA; 46537 MW; EBFDDBD50722DD CRC64;

Query Match 20.5%; Score 426.5; DB 1; Length 417;
 Best Local Similarity 31.0%; Pred. No. 8.8e-21;
 Matches 107; Conservative 64; Mismatches 119; Indels 55; Gaps 10;

44 YALIFVGVGNLVCVVRHOTLKT-PTNYLFLSAVSDLVLLGMPLTEYV-MHN 101
 Db YSLIFLGTAGNALSVHVKATGRGRRLRYHVLALSALLLLLSYVMELYNFWSH 98
 102 YPLFLFPGVGC---YFKTALFEVCAFSILSVTVSEVERVAIVHFFRAKLESTRRLARI 158
 Db YPLVFGDLGGRGYF---VRELCAVATVLSVLSAERCAVCOPLARARLLTPRRCRRL 155
 159 LSLVMSSESVFSLPNTSIIKIKQ-HEPNCSSVPGSATCTV----- 198
 Db LSLVMSSESVFSLPNTSIIKIKQ-HEPNCSSVPGSATCTV----- 198
 156 LSLVMSSESVFSLPNTSIIKIKQ-HEPNCSSVPGSATCTV----- 198
 Db LSLVMSSESVFSLPNTSIIKIKQ-HEPNCSSVPGSATCTV----- 198
 199 TRMNVYNNIIQATSF-----LFYILPMT-----LSIVLYLMLRL 235
 Db TRMNVYNNIIQATSF-----LFYILPMT-----LSIVLYLMLRL 235

Db 216 RSPLMELTAILNGITVNHVALYSQVPSASAOVNSIPSRLLELSEGLGFTWRKTL 275
 Qy 236 KRDESLFANKVAVINHRPSKSTYKMLFVLVFAICWPFPHDRLEFSEVE--WTESL 293
 Db 276 GVOASLVRHKDASOIR--SLQHSQVLRALVAVYVICWLPYHARRLWCYIPDDGWTDEL 333
 Qy 294 AAVENLHVYSGVEFYLSAANPIIYNLLSRFRAERFNVYSPYC 338
 Db 334 YDFHYHYVMYNTLFEYVSAVTPVLYNAVSSPFKLLFLESLSLJC 378

RESULT 11
 TRFR_HUMAN
 ID TRFR_HUMAN STANDARD; PRT; 398 AA.
 AC P34981;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE THYROTROPIN-RELEASING HORMONE RECEPTOR (TRH-R) (THYROLIBERIN
 DE RECEPTOR).
 GN TRHR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93371401; PubMed-8395824;
 RA Matre V., Karlens H.E., Wright M.S., Lundell I., Fieidheim K.,
 RA Gabrielsen O.S., Larhammer D., Gautvik K.M.;
 RT "Molecular cloning of a functional human thyrotropin-releasing
 RT hormone receptor.";
 RL Biochem. Biophys. Res. Commun. 195:179-185(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93384596; PubMed-8396925;
 RA Yamada M., Monden T., Satoh T., Satoh N., Murakami M., Iriuchijima T.,
 RA Kakegawa T., Mori M.;
 RT "Pituitary adenomas of patients with acromegaly express thyrotropin-
 RT releasing hormone receptor messenger RNA: cloning and functional
 RT expression of the human thyrotropin-releasing hormone receptor
 RT gene.";
 RL Biochem. Biophys. Res. Commun. 195:737-745(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94063224; PubMed-8243797;
 RA Duthie S.M., Taylor P.L., Anderson L., Cook J., Eldne K.A.;
 RT "Cloning and functional characterisation of the human TRH receptor.";
 RL Mol. Cell. Endocrinol. 95:R11-R15(1993).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95002135; PubMed-7918619;
 RA Hinuma S., Hosoya M., Ogi K., Tanaka H., Nagai Y., Onda H.;
 RT "Molecular cloning and functional expression of a human thyrotropin-
 RT releasing hormone (TRH) receptor gene.";
 RL Biochim. Biophys. Acta 1219:251-259(1994).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-96355621; PubMed-8703031;
 RA Iwasaki T., Yamada M., Satoh T., Konaka S., Ren Y., Hashimoto K.,
 RA Kohga H., Kato Y., Mori M.;
 RT "Genomic organization and promoter function of the human thyrotropin-
 RT releasing hormone receptor gene.";
 RL J. Biol. Chem. 271:22183-22188(1996).
 RN [6]
 RP SEQUENCE OF 1-263 FROM N.A.
 RC TISSUE-Placenta;
 RX MEDLINE-99101143; PubMed-9886052;
 RA Matre V., Hovring P.I., Orstavik S., Frengen E., Rlan E.,
 RA Velickovic Z., Murray-Mcintosh R.P., Gautvik K.M.;
 RT "Structural and functional organization of the gene encoding the
 RT human thyrotropin-releasing hormone receptor.";

```

RL J. Neurochem. 72:40-50(1999).
CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
CC PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D16845; BAA04120.1; -.
DR EMBL: X75071; CA52965.1; -.
DR EMBL: X72089; CA50979.1; -.
DR EMBL: S75283; AAB32222.1; -.
DR EMBL: S75281; AAB32222.1; JOINED.
DR EMBL: D85376; BAA12796.1; -.
DR EMBL: D85375; BAA12796.1; JOINED.
DR EMBL: A1011701; CA09746.1; -.
DR PIR: JN0759; JN0759.
DR PIR: S40682; S40682.
DR PIR: JN0708; JN0708.
DR GCRDB: GCR_0765; -.
DR GCRDB: GCR_0799; -.
DR GCRDB: GCR_0880; -.
DR GCRDB: GCR_1111; -.
DR GCRDB: GCR_1148; -.
DR MIM: 188545; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PRINTS: PR00751; THYROLIBRIN.
DR PROSITE: PS00237; G-PROTEIN_RECIP_F1_1; 1.
DR PROSITE: PS0262; G-PROTEIN_RECIP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 28
FT TRANSMEM 29 51
FT DOMAIN 52 61
FT TRANSMEM 62 83
FT DOMAIN 84 99
FT TRANSMEM 100 121
FT DOMAIN 122 144
FT TRANSMEM 145 168
FT DOMAIN 169 193
FT TRANSMEM 194 215
FT DOMAIN 216 266
FT TRANSMEM 267 288
FT DOMAIN 289 296
FT TRANSMEM 297 319
FT DOMAIN 320 398
FT CARBOHD 3 3
FT CARBOHD 10 10
FT SEQUENCE 398 AA; 45084 MW; FE920B5FE293D3E CRC64;
SQ

```

Query Match 20.0%; Score 416; DB 1; Length 398;
 Best Local Similarity 31.5%; Pred. No. 4e-20;
 Matches 114; Conservative 66; Mismatches 120; Indels 62; Gaps 11;

```

DB 132 KAOFICFSRANKIIFVAFSTLYCMLFLLDLINISTYKDAIY----SCGYKISRMV 187
OY 205 YNLIIGATSEFLYILPNTLSVLYLMLGLR-----KRDESLANKVAV 248
DB 188 YSPIYIMDGVYVPMILATVLYGFIARILELNPISDPKNSKTKWKNDSHTQNTLVN 247
OY 249 NHR-----PSRKSVTMLFVLYVFAICWTPFH---VDRLFES--FVEEWTESLAIV 296
DB 248 NNSNRCFNSTVSRSKQVTMLAVVLYFLMLMPPYTLTVVNSFLSSPQENW----- 300
OY 297 ENLIHVSGVFYFLSSAVNPDIYNLSRFRFAFRNVSPTCKWCHPRHPOGPAOKII 356
DB 301 ---FLFCRCICILNSAINPVIYNLSQKRAFR-----KLCNCKOKPTEKPAHYSV 350
OY 357 FL 358
DB 351 AL 352

RESULT 12
TRFR_BOVIN STANDARD: PRT: 398 AA.
ID TRFR_BOVIN
AC 046539;
DT 20-AUG-2001 (Rel. 40; Created)
DT 20-AUG-2001 (Rel. 40; Last sequence update)
DT 20-AUG-2001 (Rel. 40; Last annotation update)
DE THYROTROPIN-RELEASING HORMONE RECEPTOR (TRH-R) (THYROLIBERIN
DE RECEPTOR).
GN TRHR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HOLSTEIN;
RX MEDLINE=98151337; PubMed=9492373;
RA Taketa M., Shimada Y., Ikeda A., Sekikawa K.;
RT "Molecular cloning of bovine thyrotropin-releasing hormone receptor
RT gene.";
RL J. Vet. Med. Sci. 60:123-127(1998).
CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
CC PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D83964; BAA24069.1; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00751; THYROLIBRIN.
DR PROSITE: PS00237; G-PROTEIN_RECIP_F1_1; 1.
DR PROSITE: PS0262; G-PROTEIN_RECIP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 28
FT TRANSMEM 29 51
FT DOMAIN 52 61
FT TRANSMEM 62 83
FT DOMAIN 84 99
FT TRANSMEM 100 121
FT DOMAIN 122 144
FT TRANSMEM 145 168
FT DOMAIN 169 193
FT TRANSMEM 194 215

```

FT DOMAIN 216 266 CYTOPLASMIC (POTENTIAL).
 FT TRAMSEM 267 288 6 (POTENTIAL).
 FT DOMAIN 269 296 EXTRACELLULAR (POTENTIAL).
 FT TRAMSEM 297 319 7 (POTENTIAL).
 FT DOMAIN 320 338 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 398 AA: 45165 MW: D9AF4B211A5701B8 CRC64:

Query Match 19.9%; Score 414; DB 1; Length 398;
 Best Local Similarity 31.1%; Pred. No. 5, 3e-20;
 Matches 114; Conservative 69; Mismatches 115; Indels 68; Gaps 12;

OY 33 RSDLSPLVSAVALIFLVY---GVMGNLLCVCIVRHQTKPTNYLSLAVSDLLVL 88
 DB 17 RAVVALEYOVYVTLVLLIIGLGIIVGINVYLVVMTKHKRTTCYLSLAVADMLVLY 76
 OY 89 -LGMP--EIVEMHNYPLFGPVGCYFETALFETVCFASLSVTTSVERVAIVHP 144
 DB 77 AAGLPMTDSIYGSW-----VGVGCLCTTYLQYLGINSSCSITAFTERIAICHP 131
 OY 145 RALSTERRALRILSLWSFSVPSLPNTSHGKIFQHPNCGSSVPGSATCTVTTPMV 204
 DB 132 KAQFLCTFSFAKKIIFVNAFTSIYCMLEFLDLNISTYKDAIVY---SCGYKISRNY 187
 OY 205 YNLIQATSFILPMTLSVLYLM-----GLRLKROESLEANKVAVN 249
 DB 188 YSTIYIMDGVFTVPPILATVLTGTLARLFLNPISDPKENSMMKNDSTHONK--N 244
 OY 250 IHR-----PSKRSVTKMLFVLVLFVFAICWTPFH--VDRLEFS--FVEEWTESL 293
 DB 245 LNSKTSNRYENSTVSSRKQYTKMLAVVILFALMLMPRYTLVYVNSLSPFQENM---- 300
 OY 294 AAVFNLIHVSGVFYLLSSAVNPITYNLSRRRAAFRRNVSPCTKCWCHRRPQCPAQ 353
 DB 301 -----FLECRICILYNSAINPVIYNLSQKFRAPR-----KLCNCKOKRPEKPN 347
 OY 354 KIIFLT 359
 DB 348 YSAVLS 353

RESULT 13
 TRER_CHICK STANDARD: PRT: 395 AA.

AC 093603;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE THYROTROPIN-RELEASING HORMONE RECEPTOR (TRH-R) (THYROLIBERIN RECEPTOR).
 DE RECEPTOR).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Plutitary;
 RX MEDLINE=98344812; PubMed=9681487;
 RA Sun Y.M., Miller R.P., Ho H., Gershengorn M.C., Illing N.;
 RT Cloning and characterization of the chicken thyrotropin-releasing hormone receptor.;
 RT Endocrinology 139:3390-3398(1998).
 CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.1sb-sdb.ch/announce/> or send an email to license1sb-sdb.ch).

DR EMBL: Y18244; CAA77091.1; -
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS: PR00237; GPCR_Rhodopsn.
 DR PRINTS: PR00751; THYROLIBERIN.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 30
 FT TRAMSEM 31 53
 FT DOMAIN 54 63
 FT TRAMSEM 64 85
 FT DOMAIN 86 101
 FT TRAMSEM 102 123
 FT DOMAIN 124 146
 FT TRAMSEM 147 170
 FT DOMAIN 171 195
 FT TRAMSEM 196 217
 FT DOMAIN 218 268
 FT TRAMSEM 269 290
 FT DOMAIN 291 298
 FT TRAMSEM 299 321
 FT DOMAIN 322 395
 FT CARBOHYD 3 3
 FT CARBOHYD 10 10
 SQ SEQUENCE 395 AA: 44697 MW: 291D9BD2718723CB CRC64:

Query Match 19.9%; Score 413.5; DB 1; Length 395;
 Best Local Similarity 30.7%; Pred. No. 5, 7e-20;
 Matches 119; Conservative 70; Mismatches 125; Indels 73; Gaps 14;

OY 42 VAVALLIFLVY---GVMGNLLCVCIVRHQTKPTNYLSLAVSDLLVL-LGMP--LE 94
 DB 29 VTLVLLIIGLGIIVGINVYLVVMTKHKRTTCYLSLAVADMLVYAGLPMTTES 88
 OY 95 IYEMHNYPLFGPVGCYFETALFETVCFASLSVTTSVERVAIVHPRAKLESTRR 154
 DB 89 LYKSW-----VGVGCLCTTYLQYLGINSSCSITAFTERIAICHPKAQFLCTFSR 143
 OY 155 ALRILSLWSFSVPSLPNTSHGKIFQHPNCGSSVPGSATCTVTTPMVYNNLIQATSE 214
 DB 144 AKRIIFVNSFASVYCMLEFLDLNIAVYKDTTVY---SCGYKRSYSPIDYMDFG 199
 OY 215 LFYILPMTLSVLYLMGRL-----KRDESLANKV-----AVNIH 251
 DB 200 IFYILPMTLSVLYLMGRL-----KRDESLANKV-----AVNIH 251
 OY 252 RPSKRSVTKMLFVLVLFVFAICWTPFH--VDRLEFS--FVEEWTESLAAVFNLIHVSGV 306
 DB 260 IASRQVTKMLAVVILFALMLMPRYTLVYVNSLSPFQENM-----FLFPCRI 309
 OY 307 EFTLSAVNPITYNLSRRRAAFRRNVSPCTKCWCHRRPQCPAQKI-----IFLTCG 361
 DB 310 CIYLSAINPVIYNLSQKFRAPRKLKN-----CHLK-RDKKPNVSAVLYNVAIKESD 363
 OY 362 HLYVELTADGQPPGSSSINTNLTRA 388
 DB 364 HFSSEIEDI-----TVNITLSSA 382

RESULT 14
 TRER_MOUSE STANDARD: PRT: 393 AA.
 AC P21761;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)

DT 01-NOV-1997 (Rel. 35, last annotation update)
 DE THYROTROPIN-RELEASING HORMONE RECEPTOR (TRH-R) (THYROLIBERIN
 RECEPTOR).
 GN TRHR.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Pituitary;
 RX MEDLINE=91088548; PubMed=2175902;
 RA Straub R.E., Frech G.C., Joho R.H., Gershengorn M.C.;
 RT "Expression cloning of a cDNA encoding the mouse pituitary
 thyrotropin-releasing hormone receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9514-9518(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92381047; PubMed=1324930;
 RA Narayanan C.S., Fujimoto J., Geras-Raake E., Gershengorn M.C.;
 RT "Regulation by thyrotropin-releasing hormone (TRH) of TRH receptor
 mRNA degradation in rat pituitary GH3 cells.";
 RL J. Biol. Chem. 267:17296-17303(1992).
 RN [3]
 RP SEQUENCE OF 332-393 FROM N.A.
 RC TISSUE-Pituitary;
 RX MEDLINE=97013702; PubMed=9156522;
 RA Jones K.E., Brubaker J.H., Chin W.W.;
 RT "An alternative splice variant of the mouse TRH receptor mRNA is the
 major form expressed in the mouse pituitary gland.";
 RL J. Mol. Endocrinol. 16:197-204(1996).
 CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
 RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
 PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib.ch)
 CC -----
 DR EMBL: M59811; AAA40480.1; -;
 DR EMBL: M94384; AAA40437.1; -;
 DR EMBL: L48936; AAA81559.1; -;
 DR PIR: A39251; A39251.
 DR GCRDB: GCR_0099; -;
 DR GCRDB: GCR_1613; -;
 DR MGD: MGI:98824; Trhr.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PRO0751; THYROLIBERIN.
 DR PROSITE: PS00237; G-PROTEIN_RECPT_FL_1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECPT_FL_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein.
 KW DOMAIN 1
 FT TRANSSEM 29 51
 FT TRANSSEM 52 61
 FT TRANSSEM 62 83
 FT TRANSSEM 84 99
 FT TRANSSEM 100 121
 FT TRANSSEM 122 144
 FT TRANSSEM 145 168
 FT TRANSSEM 169 193
 FT TRANSSEM 194 215
 FT TRANSSEM 216 266
 FT TRANSSEM 267 288
 FT TRANSSEM 289 296
 FT TRANSSEM 297 319
 FT DOMAIN 320 393
 CC CYTOPLASMIC (POTENTIAL).

FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 393 AA; 44559 MW; 8739875D1A0FC6C CRC64;
 Query Match 19.8%; Score 412; DB 1; Length 393;
 Best Local Similarity 29.8%; Pred. No. 7,1e-20;
 Matches 114; Conservative 74; Mismatches 114; Indels 80; Gaps 13;
 QY 40 VSAVALIFL-----VGMGNLLVCMVIVRHOTLTPINYLFLSLAVSDLVL-LG 90
 DB 20 VALEQVAVTILLVIIICIGIGINIMVLLVMRTKMRPTNCYLSLAVDLMLVLAG 79
 QY 91 MPL---EYEMHNHNPFLPGVGCYEKTALFETVCFASLTSTVTSVERVYAVIPRAK 147
 DB 80 LPINTDSIYGSW---YGVGCGCTIYLOYLGIMASCSSTFAETIEVIAICHPAQ 134
 QY 148 LESTRRRLRLILSVSFSVPSLEPNTSIHGIFQHPFGSSVPSGATCTYTKPMVYNL 207
 DB 135 FLCTFSRAKKIIFVWATFSYCMFLDLINISTYKANAVV---SCGKISRNYS 190
 QY 208 IIOATSPFLYILPMLISVLYLM-----GLRLKDESLKANKAVNINR 252
 DB 191 IYLMDFGVYVPMILATLVYGFIRILFLNPISDPKENSIMKRNDSIHQNK-MLNINA 249
 QY 253 -----PSKRSVTKMLFVLVFAICMPFH---VDRLFEFS--FVEEWTESLAAVFN 298
 DB 250 TNRCNSVTSVSSKQVTKMLAVVILFALMLMPYRFLVAVNSFLSSPFCNM----- 300
 QY 299 LIHVSGVEFYSSAVNPDIYMLSRFAAFRNVSPYCKCHRRHPOGPAOKII-- 356
 DB 301 -FLFRCRICIYNSAINPIYVILMSQKFRANR-----KICNCKQKPTKANAAYVAL 352
 QY 357 -----FLTECHLVELTE 368
 DB 353 NYSVIKESDPFSTLEDITVTD 374
 RESULT 15
 TRFR_SHEEP STANDARD: PRT; 398 AA.
 AC Q28596;
 DT 01-NOV-1997 (Rel. 35, created)
 DT 01-NOV-1997 (Rel. 35, last sequence update)
 DT 15-JUL-1998 (Rel. 36, last annotation update)
 DE THYROTROPIN-RELEASING HORMONE RECEPTOR (TRH-R) (THYROLIBERIN
 RECEPTOR).
 GN TRHR.
 OS Ovis aries (Sheep).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Caprinae; Ovis.
 CC NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Pituitary;
 RX MEDLINE=97200773; PubMed=9048604;
 RA Bockmann J., Boeckers T.M., Winter C., Wiltkowski W., Winterhoff H.,
 RA Deufel T., Kreutz M.R.;
 RT "Thyrotropin expression in hypophyseal pars tuberalis-specific cells
 is 3,5,3'-triiodothyronine, thyrotropin-releasing hormone, and pit-1
 independent.";
 RL Endocrinology 138:1019-1028(1997).
 CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
 RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
 PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X95285; CMA64606.1; -
DR GCRDB; GCR_1294; -
DR InterPro; IPRO00276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00751; THYROLIBRIN.
DR PROSITE; PS00237; G_PROTEIN_RECIP_FL_2; 1.
DR PROSITE; PS50262; G_PROTEIN_RECIP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 28 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 29 51 1 (POTENTIAL).
FT DOMAIN 52 61 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 62 83 2 (POTENTIAL).
FT DOMAIN 84 99 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 100 121 3 (POTENTIAL).
FT DOMAIN 122 144 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 145 168 4 (POTENTIAL).
FT DOMAIN 169 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215 5 (POTENTIAL).
FT DOMAIN 216 266 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 267 288 6 (POTENTIAL).
FT DOMAIN 289 296 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 297 319 7 (POTENTIAL).
FT DOMAIN 320 398 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 398 AA: 45088 MW: 375A311D3DD2A61A CRC64;

Query Match 19.8%; Score 412; DB 1; Length 398;
Best Local Similarity 31.5%; Pred No. 7.2e-20;
Matches 115; Conservative 67; Mismatches 115; Indels 68; Gaps 12;

OY 33 RSDSLPVSVAVALFLV---GVMGNLLVCWVIVRQTLKPTNYLFLSLAVSDLLVL 88
DB 17 RAVVALEYQVITLLVLLICGLIYINIVLVVMTKMRPTNCLYLSLAVADLMVLV 76
OY 89 -LGMP-EEYEMHNYFLFGPGCYFKTALFETVCASISVTVSVERVALVHPF 144
DB 77 AAGLPRTIDSIYGSW---VYGVGCLCTITLYLYGAINASSCSITAFIERIALCHPI 131
OY 145 RAKLESTRRLRLIISLVMSFSVFSLPNTSINGIKFQHFPGSSVPGSATCTVKPMV 204
DB 132 KQOICTFSRAKKIIFVAFISYICMLWFLDLNISTYKDAIV---SCGYKISRNY 187
OY 205 VNLITQATSFLFYLLPMLISLVLYLMGLR-----KRDESLANKRAVN 249
DB 188 YSPIYLMDFGVVPMILATVLYGFIARILFSPIDPKENSNTWKNDSTHONK---N 244
OY 250 IHR-----PSRKSVTKMLFLVLYVFAICWTPFH---VDRLEFS--FVEEWTESL 293
DB 245 LMSKTSNRKFNSTVSRKQVTKMLAVVTLFALLMMPYRFLVAVNSFLSPQENW--- 300
OY 294 AAVFNLIHVSGVFYFLSSAVNPDIYNLISRRRAAFRNVAVSPTCKWCHPRHPOGPPAO 353
DB 301 -----FLLFRCICILYLSAINPVIYNLMSQKRAFR-----KLCNCKOKPYEKPRAN 347
OY 354 KTIIFL 358
DB 348 YSVAL 352

Search completed: April 22, 2002, 14:19:42
Job time: 145 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: Apr11 22, 2002, 14:18:12 ; Search time 21.56 Seconds
(without alignments)
412.282 Million cell updates/sec

Title: US-09-609-146-25
Perfect score: 2076
Sequence: 1 MGRLENSWIDHPLMKYILNS.....GQSSIHNTLTAPCAGEVP 395

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCITUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	465.5	22.4	364	4	US-09-077-675A-16 Sequence 16, Appl
2	465	22.4	353	4	US-09-077-675A-3 Sequence 3, Appl
3	463.5	22.3	361	4	US-09-077-675A-8 Sequence 8, Appl
4	463.5	22.3	366	4	US-09-077-675A-13 Sequence 13, Appl
5	437	21.1	410	3	US-08-858-876A-2 Sequence 2, Appl
6	437	21.1	410	4	US-09-472-880-2 Sequence 2, Appl
7	433.5	20.9	353	1	US-08-118-270-45 Sequence 45, Appl
8	433.5	20.9	353	5	PCT-US93-08528-45 Sequence 45, Appl
9	430.5	20.7	416	4	US-08-858-876A-4 Sequence 4, Appl
10	430.5	20.7	416	3	US-09-472-880-4 Sequence 4, Appl
11	423	20.4	302	4	US-09-077-675A-2 Sequence 2, Appl
12	421.5	20.3	302	4	US-09-077-675A-7 Sequence 7, Appl
13	416	20.0	398	4	US-08-288-663A-1 Sequence 1, Appl
14	412	19.8	393	1	US-07-629-1041-3 Sequence 3, Appl
15	400.5	19.3	400	3	US-08-889-108-8 Sequence 8, Appl
16	400.5	19.3	400	5	PCT-US94-10358-8 Sequence 8, Appl
17	399.5	19.2	384	4	US-08-103-170-10 Sequence 10, Appl
18	395.5	19.1	400	4	US-08-188-275A-2 Sequence 2, Appl
19	392	18.9	341	1	US-08-118-270-48 Sequence 48, Appl
20	392	18.9	341	5	PCT-US93-08528-48 Sequence 48, Appl
21	389.5	18.8	319	4	US-08-832-399-2 Sequence 2, Appl
22	389.5	18.8	319	4	US-09-372-498-2 Sequence 2, Appl
23	388.5	18.7	398	2	US-08-288-663A-15 Sequence 15, Appl
24	388.5	18.3	411	1	US-07-937-609-21 Sequence 21, Appl
25	380.5	18.3	411	4	US-08-029-170-21 Sequence 3, Appl
26	380	18.3	391	2	US-08-454-549-3 Sequence 3, Appl
27	380	18.3	391	3	US-08-454-552-3 Sequence 3, Appl

28	380	18.3	398	1	US-08-149-093A-5 Sequence 5, Appl
29	380	18.3	398	2	US-08-911-245-5 Sequence 5, Appl
30	380	18.3	398	3	US-08-889-108-2 Sequence 2, Appl
31	380	18.3	398	4	US-08-120-601B-2 Sequence 2, Appl
32	380	18.3	398	4	US-08-188-275A-3 Sequence 3, Appl
33	380	18.3	398	4	US-08-387-707-16 Sequence 16, Appl
34	380	18.3	398	4	US-09-510-473-5 Sequence 5, Appl
35	380	18.3	398	5	PCT-US94-10358-2 Sequence 2, Appl
36	374	18.0	391	3	US-08-676-351-4 Sequence 4, Appl
37	374	18.0	398	3	US-09-170-331-5 Sequence 5, Appl
38	373	18.0	387	1	US-08-196-989B-14 Sequence 14, Appl
39	373	18.0	387	2	US-08-760-936-14 Sequence 10, Appl
40	371	17.9	451	4	US-08-430-286A-10 Sequence 2, Appl
41	370	17.8	356	4	US-08-430-286A-2 Sequence 19, Appl
42	368	17.7	369	1	US-07-937-609-19 Sequence 19, Appl
43	368	17.7	369	4	US-08-029-170-19 Sequence 19, Appl
44	366.5	17.7	465	4	US-08-090-369-1 Sequence 1, Appl
45	363.5	17.5	384	3	US-09-071-434-3 Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-077-675A-16
Sequence 16, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077, 675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ. ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 364 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-077-675A-16
Query Match 22.4%; Score 465.5; DB 4; Length 364;

Best Local Similarity 34.6%; Pred. No. 5.2e-31;
Matches 104; Conservative 65; Mismatches 107; Indels 25; Gaps 6;

QY 40 VSAVYALIFLVGVGNLJYCMVIVRHOTLKPTNYVLFSLAVSDLLVLLGMPLEIEM 99
Db 45 VTATCVALLFVGVGNLJYCMVIVRHOTLKPTNYVLFSLAVSDLLVLLGMPLEIEM 103
QY 100 HNPFLGPGVCYKTLAFETVCFASILSVTVSVERYVAIVHFRKLESTRRARIL 159
Db 104 QYRWMNGLNCKLFGQVSESCYATVLTITALSVERYPALCFPLRAKVVYTKRVLVI 163
QY 160 SLWSESVESLPMTSTHGKIFQHFPGSSVPGSATCTVTK-----PMVYLLII 209
Db 164 LVIAVAFCSGPIFLVGVGEH---NGTDPDRTNECRATEFAVRSGLLTVMVV----- 215
QY 210 QATSEFLYIIPMTLISVLYLMGLRLKRDSELEANKVAVNIHRPSKSVTKMLFVLVLF 269
Db 216 ---SVFFFLPVFCITVLYSLIGRKLRRRGDA--VGASLRDQNHQYKMLAVVFAF 270
QY 270 AICWTPHYDRLFPS-EVEWTESTLAVFNLIVHVGCVFFYLSAVNPDIYNLSRRERA 328
Db 271 ILCWLPYHGVGRYLSKSPGSELEIAQISQCNVSVFLYLSAANPILYNIMSKRYR 330
QY 329 A 329
Db 331 A 331

RESULT 2
US-09-077-675A-3
Sequence 3, Application US/09077675A
Patent No. 6242199

GENERAL INFORMATION:

APPLICANT: Pal, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077, 675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Cocuzzo, Anna L.

REGISTRATION NUMBER: 42,452

REFERENCE/DOCKET NUMBER: 19590P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 732-594-1273

TELEFAX: 732-594-4720

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 353 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
US-09-077-675A-3

Query Match 22.4%; Score 465; DB 4; Length 353;
Best Local Similarity 34.2%; Pred. No. 5.5e-31;
Matches 103; Conservative 68; Mismatches 106; Indels 24; Gaps 6;

QY 40 VSAVYALIFLVGVGNLJYCMVIVRHOTLKPTNYVLFSLAVSDLLVLLGMPLEIEM 99
Db 33 VTATCVALLFVGVGNLJYCMVIVRHOTLKPTNYVLFSLAVSDLLVLLGMPLEIEM 91
QY 100 HNPFLGPGVCYKTLAFETVCFASILSVTVSVERYVAIVHFRKLESTRRARIL 159
Db 92 QYRWMNGLNCKLFGQVSESCYATVLTITALSVERYPALCFPLRAKVVYTKRVLVI 151
QY 160 SLWSESVESLPMTSTHGKIFQHFPGSSVPGSATCTVTK-----PMVYLLII 209
Db 152 LVIAVAFCSGPIFLVGVGEH---NGTDPDRTNECRATEFAVRSGLLTVMVV----- 203
QY 210 QATSEFLYIIPMTLISVLYLMGLRLKRDSELEANKVAVNIHRPSKSVTKMLFVLVLF 269
Db 204 ---SVFFFLPVFCITVLYSLIGRKLRRRGDA--VGASLRDQNHQYKMLAVVFAF 259
QY 270 AICWTPHYDRLFPS-EVEWTESTLAVFNLIVHVGCVFFYLSAVNPDIYNLSRRERA 328
Db 260 ILCWLPYHGVGRYLSKSPGSELEIAQISQCNVSVFLYLSAANPILYNIMSKRYR 319
QY 329 A 329
Db 320 A 320

RESULT 3
US-09-077-675A-8
Sequence 8, Application US/09077675A
Patent No. 6242199

GENERAL INFORMATION:

APPLICANT: Pal, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077, 675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Cocuzzo, Anna L.

REGISTRATION NUMBER: 42,452

REFERENCE/DOCKET NUMBER: 19590P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 732-594-1273

TELEFAX: 732-594-4720

TELEX:

INFORMATION FOR SEQ ID NO: 8:


```

: SEQUENCE CHARACTERISTICS:
:   LENGTH: 361 amino acids
:   TYPE: amino acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
:   MOLECULE TYPE: protein
us-09-077-675A-8

```

```

Query Match      22.3%: Score 463.5; DB 4; Length 361;
Best Local Similarity 34.9%: Pred. No. 7.6e-31;
Matches 106; Conservative 66; Mismatches 107; Indels 25; Gaps 7;

```

```

Oy 40 VSAVALIFLVGMGNLLVCMVIVRHQTLKPTNTYYLSLAVSDLLVLLGMPLEIYEMW 99
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 41 VTATCVALLFVVGIGACNLMLVLSFRRLRTTNLYLSMAFSDLLIFLC-MPLDLVRLM 99
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 100 HNPFLFGVGCYFETALFETVCFASILSVTVSVERYVAIVHPRAKLESTRRALRL 159
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 100 QYRPNFGDLCKLFQFVSESCYATVLTITLSEYRFAICFPLRAKVVYTKGKVKLYI 159
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 160 SLVMSFSVFSLPNTHIGIKQHFHPNCGSSVPGSATC-----IVTKRMVYNLI 209
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 160 FVIMAVAFCSAGPIFVLGVGEHE--NGTDPMDTNECRPTEFAVBSGLTVVWV----- 211
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 210 QATSFLEYILPMTLSVLYLMLGRKDESLANKVAVNIHRPSKSVTKMLFVLVLF 269
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 212 ---SIFFLPVFCULTVLYSLGRKLMRRRGDA--VCGASLDQHKQYVKMLAVVFAF 267
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 270 AICWTPFHVDRLFES-EVEEMTESLAAVENLIHVVGVEFYLSAVNPITYNLSRRRR- 327
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 268 ILCMLPFHVGRYLFESKSEPGSLLEIAQISQYCNLVSFLYLSAAMINPILYNIMSKKRYV 327
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 328 AAFR 331
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 328 AVFR 331
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 4
US-09-077-675A-13
: Sequence 13, Application US/09077675A
: Patent No. 6242199
: GENERAL INFORMATION:
:   APPLICANT: Pal, Lee-Yuh
:   APPLICANT: Feighner, Scott C.
:   APPLICANT: Howard, Andrew D.
:   APPLICANT: Pong, Sheng-Shung
:   APPLICANT: Van Der Ploeg, Leonardus H.T.
:   TITLE OF INVENTION: RECEPTOR ASSAY
:   NUMBER OF SEQUENCES: 16
:   CORRESPONDENCE ADDRESS:
:     ADDRESSEE: Merck & Co., Inc.
:     STREET: P.O. Box 2000, 126 E. Lincoln Ave.
:     CITY: Rahway
:     STATE: NJ
:     COUNTRY: USA
:     ZIP: 07065-0900
:   COMPUTER READABLE FORM:
:     MEDIUM TYPE: Diskette
:     COMPUTER: IBM Compatible
:     OPERATING SYSTEM: DOS
:     SOFTWARE: FASTSEQ for Windows Version 2.0
:   CURRENT APPLICATION DATA:
:     APPLICATION NUMBER: US/09/077.675A
:     FILING DATE: 3-JUN-1998
:   CLASSIFICATION:
:     PRIOR APPLICATION DATA:
:       APPLICATION NUMBER:
:       FILING DATE:
:       ATTORNEY/AGENT INFORMATION:
:         NAME: Cocuzzo, Anna L.
:         REGISTRATION NUMBER: 42,452
:         REFERENCE/DOCKET NUMBER: 19590P

```

```

: TELECOMMUNICATION INFORMATION:
:   TELEPHONE: 732-594-1273
:   TELEFAX: 732-594-4720
:   TELEX:
:   INFORMATION FOR SEQ ID NO: 13:
:     SEQUENCE CHARACTERISTICS:
:       LENGTH: 366 amino acids
:       TYPE: amino acid
:       STRANDEDNESS: single
:       TOPOLOGY: linear
:       MOLECULE TYPE: protein
us-09-077-675A-13

```

```

Query Match      22.3%: Score 463.5; DB 4; Length 366;
Best Local Similarity 34.9%: Pred. No. 7.7e-31;
Matches 106; Conservative 66; Mismatches 107; Indels 25; Gaps 7;

```

```

Oy 40 VSAVALIFLVGMGNLLVCMVIVRHQTLKPTNTYYLSLAVSDLLVLLGMPLEIYEMW 99
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 46 VTATCVALLFVVGIGACNLMLVLSFRRLRTTNLYLSMAFSDLLIFLC-MPLDLVRLM 104
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 100 HNPFLFGVGCYFETALFETVCFASILSVTVSVERYVAIVHPRAKLESTRRALRL 159
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 105 QYRPNFGDLCKLFQFVSESCYATVLTITLSEYRFAICFPLRAKVVYTKGKVKLYI 164
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 160 SLVMSFSVFSLPNTHIGIKQHFHPNCGSSVPGSATC-----IVTKRMVYNLI 209
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 165 FVIMAVAFCSAGPIFVLGVGEHE--NGTDPMDTNECRPTEFAVBSGLTVVWV----- 216
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 210 QATSFLEYILPMTLSVLYLMLGRKDESLANKVAVNIHRPSKSVTKMLFVLVLF 269
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 217 ---SIFFLPVFCULTVLYSLGRKLMRRRGDA--VCGASLDQHKQYVKMLAVVFAF 272
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 270 AICWTPFHVDRLFES-EVEEMTESLAAVENLIHVVGVEFYLSAVNPITYNLSRRRR- 327
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 273 ILCMLPFHVGRYLFESKSEPGSLLEIAQISQYCNLVSFLYLSAAMINPILYNIMSKKRYV 332
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 328 AAFR 331
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 328 AVFR 336
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 5
US-08-858-876A-2
: Sequence 2, Application US/08858876A
: Patent No. 6022856
: GENERAL INFORMATION:
:   APPLICANT: Daniel CAPUT
:   APPLICANT: Pascale CHALON
:   APPLICANT: Pascale FERRARA
:   APPLICANT: VITA NATALIO
:   TITLE OF INVENTION: Type 2 Neurotensin Receptor
:   NUMBER OF SEQUENCES: 12
:   CORRESPONDENCE ADDRESS:
:     ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
:     STREET: 400 Seventh Street
:     CITY: Washington
:     STATE: D.C.
:     COUNTRY: USA
:     ZIP: 20004
:   COMPUTER READABLE FORM:
:     MEDIUM TYPE: Floppy disk
:     COMPUTER: IBM PC compatible
:     OPERATING SYSTEM: PC-DOS/MS-DOS
:     SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
:   CURRENT APPLICATION DATA:
:     APPLICATION NUMBER: US/08/858,876A
:     FILING DATE: 19-SEP-1997
:   CLASSIFICATION: 536
:   PRIOR APPLICATION DATA:
:     APPLICATION NUMBER: PCT/FR 9723204

```

ATTORNEY/AGENT INFORMATION:
NAME: Player, William E.
REGISTRATION NUMBER: 31,049
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 410 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-858-876A-2

Query Match 21.1%; Score 437; DB 3; Length 410;
Best Local Similarity 30.2%; Pred. No. 1.3e-28;
Matches 108; Conservative 73; Mismatches 123; Indels 54; Gaps 11;

QY 44 YALIFLVGNGNLVCMVYVRHQT-LT-PTNYLFSLANSDDLVLGMPLEIYE-MWHN 101
DB 39 YALIMALGAAGNALSVHVVYLKARAGRGLRHVHLSLALAGLLLVGPVELYSEVWFH 98
QY 102 YPFLFPGVGC---YFKTALFETVCFAISLVTVTSEYVVAIVHPRAKLESTRRALRT 158
DB 99 YPWFEDLCGRGYF---VHELCAVATVLSVAGLSAEKCLAVCQPLRAKSLTPRRTRWL 155
QY 159 LSLWSESVVFSLPNTSINGIKFQ-HFPNGSSVPGSATCTVTKPMWVYMLIQATSEFLFY 217
DB 156 VALSMASISGLALPMVAVINGOKHELTADGEPPASRVCTVLVSRALQYFIQVNVLYSF 215
QY 218 ILPMTLISVL-----YLMG-----LRKDESL-----EA 243
DB 216 VLPLATLAFNGVTVSHLALCSQVPTSTPGSSPTSRLELLSEGLLSFIYVKKTFIOG 275
QY 244 NKAVAVIHPRSRK-----SVTKMLFVLVYFAICWTPFHVDRLFFESFV---EEMTESLA 296
DB 276 GOVSLVRHKDVRIRSLQSRVQVLRALIVMVCMLPYHARRLMTCYVDDAWTIDPLYNF 335
QY 297 FNLIHVSGVFYLLSSAVNPITVYLLSRFRRAFRNVVSPCKMCHPRNR----PGCP 350
DB 336 YHFFVMTNTLFYVSSAVTPPLLYNAVSSSRKFLFLEAVSSLCGEHHPMKRLPPKQSP 393

RESULT 6
US-09-472-880-2
Sequence 2, Application US/09472880
Patent No. 6274333
GENERAL INFORMATION:
APPLICANT: Daniel CAPUT
PASCAL CHALON
PASCAL FERRARA
VITA NATALIO
TITLE OF INVENTION: Type 2 Neurotensin Receptor
(hmr-R2)
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
STREET: 400 Seventh Street
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/472,880
FILING DATE: 28-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 9723204
FILING DATE: 17-MAR-1997

ATTORNEY/AGENT INFORMATION:
NAME: Player, William E.
REGISTRATION NUMBER: 31,049
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 410 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-472-880-2

Query Match 21.1%; Score 437; DB 4; Length 410;
Best Local Similarity 30.2%; Pred. No. 1.3e-28;
Matches 108; Conservative 73; Mismatches 123; Indels 54; Gaps 11;

QY 44 YALIFLVGNGNLVCMVYVRHQT-LT-PTNYLFSLANSDDLVLGMPLEIYE-MWHN 101
DB 39 YALIMALGAAGNALSVHVVYLKARAGRGLRHVHLSLALAGLLLVGPVELYSEVWFH 98
QY 102 YPFLFPGVGC---YFKTALFETVCFAISLVTVTSEYVVAIVHPRAKLESTRRALRT 158
DB 99 YPWFEDLCGRGYF---VHELCAVATVLSVAGLSAEKCLAVCQPLRAKSLTPRRTRWL 155
QY 159 LSLWSESVVFSLPNTSINGIKFQ-HFPNGSSVPGSATCTVTKPMWVYMLIQATSEFLFY 217
DB 156 VALSMASISGLALPMVAVINGOKHELTADGEPPASRVCTVLVSRALQYFIQVNVLYSF 215
QY 218 ILPMTLISVL-----YLMG-----LRKDESL-----EA 243
DB 216 VLPLATLAFNGVTVSHLALCSQVPTSTPGSSPTSRLELLSEGLLSFIYVKKTFIOG 275
QY 244 NKAVAVIHPRSRK-----SVTKMLFVLVYFAICWTPFHVDRLFFESFV---EEMTESLA 296
DB 276 GOVSLVRHKDVRIRSLQSRVQVLRALIVMVCMLPYHARRLMTCYVDDAWTIDPLYNF 335
QY 297 FNLIHVSGVFYLLSSAVNPITVYLLSRFRRAFRNVVSPCKMCHPRNR----PGCP 350
DB 336 YHFFVMTNTLFYVSSAVTPPLLYNAVSSSRKFLFLEAVSSLCGEHHPMKRLPPKQSP 393

RESULT 7
US-08-118-270-45
Sequence 45, Application US/08118270
Patent No. 5508384
GENERAL INFORMATION:
APPLICANT: Murphy, Randall B.
APPLICANT: Schuster, David I.
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NETMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,270
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033

; FILING DATE: 3-JUN-1
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:

```

: APPLICATION NUMBER: 198309/1993
: FILING DATE: 10-AUG-1993
: APPLICATION NUMBER: 286986/1993
: FILING DATE: 16-NOV-1993
: APPLICATION NUMBER: 325215/1993
: FILING DATE: 22-DEC-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Resnick, David S
: REGISTRATION NUMBER: 34,235
: REFERENCE/DOCKET NUMBER: 44612
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 398 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-288-663A-1

```

```

Query Match          20.0%; Score 416; DB 2; Length 398;
Best Local Similarity 31.5%; Pred. No. 6,8e-27;
Matches 114; Conservative 66; Mismatches 120; Indels 62; Gaps 11;

```

```

QY 33 RSDLSLPVSAVALILFLY-----GVMGNLVCMYIVRHQTLKTPPNYYLSLAVSDLLVL 88
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 17 RAVALLEYQVVTLLVILICGLIGVGNIMVLLVMRTKHKRTPTNCYLSLAVADLVAVL 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 89 -LCMPL---EYEMAHNYPLFCGVCYFETALFETVCFASISLVTVSERYVAIVHPF 144
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 AAGLPNTDSIGSM-----VGVGCLCTITLYOYGINASSCSITAFTERIATICHPI 131
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 145 RALESTRRALRLISLWMSFVSFLPNTSINGIKFQHPNGSSVPGSATCTVTRPMVY 204
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 132 KAQFLCFESRAKKIIFVMAFTSLYCMLEFLLDINSTYKDAIVI-----SCGYKISRNY 187
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 205 YNLIQATSFYILPMTLSLVLYLMGLRL-----KRDESLANKVAV 248
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 YSPYLMDFGVYVPMILATVYGFARILFLNPISDPKSKMKNDSIQHNTNLN 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 249 NIHR-----PSRKSVTKMLFLVLYVFAICWTPFH---VDRLFFS--FVEEMTESLA 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 NTRNCSTVSSRKQVTKMLAVVILFALLMMPYRTLVVNSLSSPQENW----- 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 297 ENLHVSGVFYLSAVNPDIYNLSRRFRAFRNVVSPCKWCHPRHRPOGPPAQKII 356
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 -FLIFCRICILNSAINPVIYNLMSQKFRFAFR-----KLCNCKOKPTEKAPANYSV 350
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 357 FL 358
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 351 AL 352

```

```

RESULT 14
US-07-629-104I-3
: Sequence 3, Application US/07629104I
: Patent No. 5288621
: GENERAL INFORMATION:
: APPLICANT: Gershengorn, Marvin C
: TITLE OF INVENTION: PITUITARY TRH RECEPTOR
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Yahwak & Associates
: STREET: 25 Skytop Drive
: CITY: Trumbull
: STATE: Connecticut
: COUNTRY: USA
: ZIP: 06611
: COMPUTER READABLE FORM:

```

```

: MEDIUM TYPE: Floppy disk
: COMPUTER: Macintosh
: OPERATING SYSTEM: MS-DOS
: SOFTWARE: Microsoft Word 3.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/629,104I
: FILING DATE: 19901214
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: George M. Yahwak
: REGISTRATION NUMBER: 26,824
: REFERENCE/DOCKET NUMBER: CRF D - 995
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (203)268-1951
: TELEFAX: (203)268-1951
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 393 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-07-629-104I-3

```

```

Query Match          19.8%; Score 412; DB 1; Length 393;
Best Local Similarity 29.8%; Pred. No. 1,4e-26;
Matches 114; Conservative 74; Mismatches 114; Indels 80; Gaps 13;

```

```

QY 40 VSAVALIFL-----GVMGNLVCMYIVRHQTLKTPPNYYLSLAVSDLLVL-LG 90
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 20 VALEYQVVTLLVILICGLIGVGNIMVLLVMRTKHKRTPTNCYLSLAVADLVAVAG 79
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 91 MPL---EYEMAHNYPLFCGVCYFETALFETVCFASISLVTVSERYVAIVHPRAK 147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 80 LPNITDSIGSM-----VGVGCLCTITLYOYGINASSCSITAFTERIATICHPIKAQ 134
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 148 LESTRRRALRLISLWMSFVSFLPNTSINGIKFQHPNGSSVPGSATCTVTRPMVY 207
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 135 FLCTFSRAKKIIFVMAFTSLYCMLEFLLDINSTYKKNVY-----SCGYKISRNYSP 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 208 TIQATSFYILPMTLSLVLYLM-----GLRLKRDSELANKVAVNIHR 252
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 191 IYLMDFGVYVPMILATVYGFARILFLNPISDPKSKMKNDSIQHNTNLN 249
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 253 -----PSRKSVTKMLFLVLYVFAICWTPFH---VDRLFFS--FVEEMTESLA 298
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 250 TNRNCSTVSSRKQVTKMLAVVILFALLMMPYRTLVVNSLSSPQENW----- 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 299 LHVSGVFYLSAVNPDIYNLSRRFRAFRNVVSPCKWCHPRHRPOGPPAQKII-- 356
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 -FLIFCRICILNSAINPVIYNLMSQKFRFAFR-----KLCNCKOKPTEKAPANYSV 352
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 357 -----FLTECHLVELTE 368
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 353 NYSVYKESDRFTELEDITVTD 374

```

```

RESULT 15
US-08-889-108-8
: Sequence 8, Application US/08889108
: Patent No. 6103492
: GENERAL INFORMATION:
: APPLICANT: Yu, Lei
: TITLE OF INVENTION: Mu Opioid Receptors: Compositions and Methods
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Arnold, White & Durkee
: STREET: P. O. Box 4433
: CITY: Houston
: STATE: TX
: COUNTRY: USA
: ZIP: 77210-4433
: COMPUTER READABLE FORM:

```

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/889,108
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/305,518
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mark B.
REGISTRATION NUMBER: 37,259
REFERENCE/DOCKET NUMBER: INDA005\WIM
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-418-3000
TELEFAX: 512-474-7577
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US:08-889-108-8

Query Match 19.3% Score 400.5; DB 3; Length 400;
Best Local Similarity 30.1%; Pred. No. 1.3e-25;

Matches 106; Conservative 66; Mismatches 133; Indels 47; Gaps 11;

QY 8 SWHDPMLKYLSTEEYLALGCPKRSDL-----SLPVSVALIFLY 50
DB 30 SWVN-----LSHLEGNLSDPCGPNRTDLGGRDSLCPPTGSPSMITAITIMALYSIVCV 83
QY 51 GVMGNLVCWVIYRHQTLKPTNYLFSLAVSDLVLLGMPLE-IYEMHNYPFLEGPV 109
C 84 GLFGNFLVWYIVRTKMTATNIYI--LALDALATST-LPQSVNYLMGTWP--FGTI 140
QY 110 GCYRTALFETVCFASISVTTSVERVAVIHPFRALKESTRRALRLISLVSFSVVF 169
DB 141 LKIVISIDVYMMFTSIFTLCTMSVDRIAVCHPVKALDFRTPRNAKLIINCNMILSSAI 200
QY 170 SLPTNSIGIKFOHFPNNSVPGSATCTVT--KPMVYNLIIOATSFLE-YLPMTLISV 226
DB 201 GLPWFMAATTYKRO-----GSIDCTLTFSSHPTWYENLLKICVFIFAFIMPLIITV 252
QY 227 LYYLMLGLRLKRDSELEANKVA-VNIHRSRKSVMKMLFVLVLFALICWTPZHVDRLFFSF 285
DB 253 CYGLMILRLKSVRLSGSKENDRLNR-----ITRMVLVVAVFIVCWTPIHIYIIRKAL 307
QY 286 VEEVTESLAAVENLIHVVSGVFYLLSSAVNPDIYNLSRRPRAFRNVVSP 337
DB 308 VTIPETTFOTVSMHFCIALG---YTNGLNPLVLAFLDENKRCRCFRECIFT 356

Search completed: April 22, 2002, 14:18:13
Job time: 326 sec

THIS PAGE BLANK (uspro)